



DR WPI: 2001-657170/75.  
XX N-PSDB; AAH43685.  
PT New variants of human AMP-activated protein kinase gamma3 subunit  
PT associated with a metabolic disease e.g. diabetes or obesity and method  
PT for determining a risk estimate of diseases in subject by detecting the  
PT variant -  
XX  
XX  
PS Disclosure: Fig 5; 25pp; English.  
XX  
CC This sequence is encoded by the full length cDNA encoding the human  
CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting  
CC the presence of the PRKAG3 DNA, or a variant, is useful in determining  
CC a risk estimate of a metabolic disease, such as diabetes or obesity,  
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon  
CC 3 variation may be a substitution of a G for a C at nucleotide 320,  
CC resulting in the amino acid substitution P71A; in exon 4 variation may  
CC be a substitution of a T for a C at nucleotide 550; and in exon 10  
CC variation may be a substitution of a T for a C at nucleotide 1037,  
CC resulting in the amino acid substitution R340W. There may also be  
CC nucleotide variation in intron 6. The numbering of these  
CC variations is based on the full length cDNA, rather than on  
CC position 1 of the open reading frame.  
XX  
XX Sequence 489 AA;  
SQ  
Query Match 100.0%; Score 2538; DB 22; Length 489;  
Best Local Similarity 100.0%; Pred. No. 7.3e-254;  
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPGLEHALRRTPSWSLGGSHQEMSPLEQENSSSNWSPAVTSSSERIRKRAKALRW 60  
DB 1 MEPGLEHALRRTPSWSLGGSHQEMSPLEQENSSSNWSPAVTSSSERIRKRAKALRW 60  
QY 61 TRQKSVEEGPPQGGEGPNSRPAABSTGLEATPKTTPLAQADPAGVGTPTPTGWDCLPSD 120  
DB 61 TRQKSVEEGPPQGGEGPNSRPAABSTGLEATPKTTPLAQADPAGVGTPTPTGWDCLPSD 120  
QY 121 CTASAAGSSTDDVELATEPPEATEWCELEGLLEERPALCLSPQAPFFKLGWDELKPKG 180  
DB 121 CTASAAGSSTDDVELATEPPEATEWCELEGLLEERPALCLSPQAPFFKLGWDELKPKG 180  
QY 181 AQIYWRMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGM 240  
DB 181 AQIYWRMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGM 240  
QY 241 LTITDFILVLRHYRSPVQIYEIQHKIETWREIYVQGCFFPLVSI SPNDSLFEAYVTL 300  
DB 241 LTITDFILVLRHYRSPVQIYEIQHKIETWREIYVQGCFFPLVSI SPNDSLFEAYVTL 300  
QY 301 IKNRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLA 360  
DB 301 IKNRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLA 360  
QY 361 VVLETAFTALTALDIFVDRVSNALPVNVECGVGLYSRFDVTHLAAQOTYNHLDMSVGEA 420  
DB 361 VVLETAFTALTALDIFVDRVSNALPVNVECGVGLYSRFDVTHLAAQOTYNHLDMSVGEA 420  
QY 421 LRQRTLCLEGLVSCQPHESLGEVIDIAREQVHRLVLDVETQHLGLGVVSLDIQALVLVS 480  
DB 421 LRQRTLCLEGLVSCQPHESLGEVIDIAREQVHRLVLDVETQHLGLGVVSLDIQALVLVS 480  
QY 481 PAGIDALGA 489  
DB 481 PAGIDALGA 489  
RESULT 2  
AAE00223  
ID AAE00223 standard; Protein; 464 AA.  
XX  
AC AAE00223;  
XX

DF 13-JUN-2001 (first entry)  
XX  
DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
XX  
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Domain 172..225  
XX /label= CBS  
XX /note= "Cystathione beta synthase domain"  
XX Misc-difference 200  
XX /note= "RN- mutation site"  
XX Domain 253..307  
XX /label= CBS  
XX /note= "Cystathione beta synthase domain"  
XX Domain 329..382  
XX /label= CBS  
XX /note= "Cystathione beta synthase domain"  
XX Domain 400..453  
XX /label= CBS  
XX /note= "Cystathione beta synthase domain"  
XX  
XX W0200120003-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 11-SEP-2000; 2000WO-EP09896.  
XX  
XX 10-SEP-1999; 99EP-0402236.  
XX 18-MAY-2000; 2000EP-0401388.  
XX  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX (ANDE/) ANDERSSON L.  
XX (LOOF/) LOOFT C.  
XX (KALM/) KALM E.  
XX  
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
XX Tannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX WPI: 2001-244810/25.  
XX N-PSDB; AAD03320.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
XX monophosphate-activated kinase for diagnosis or treatment of disorders  
XX associated with energy metabolism such as diabetes, obesity, and  
XX myopathy -  
XX  
XX Claim 5; Fig 3; 71pp; English.  
XX  
XX The present sequence is human adenosine monophosphate  
XX (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
XX complete PRKAG3. Mutation in PRKAG3 results in an altered regulation of  
XX carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
XX useful as therapeutic for treating carbohydrate metabolism disorders such  
XX as diabetes, obesity, and disorders associated with muscle metabolism  
XX such as myopathy and cardiovascular diseases, to modulate AMPK  
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence  
XX and its functionally altered mutants are useful for the diagnostic  
XX evaluation, genetic testing and prognosis of a metabolic disorder,  
XX preferably a carbohydrate metabolism disorder. Primers that can detect  
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
XX useful for detecting a dysfunction of carbohydrate metabolism resulting  
XX from the expression of a functionally altered allele of PRKAG3.  
XX transgenic animal and host cell transformed with PRKAG3 or a  
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
XX screening compounds able to modulate AMPK activity. Nucleic acid  
XX encoding PRKAG3 is useful for detecting mutations in a PRKAG3 gene, or  
XX in a sequence encoding the first cystathione beta synthase (CBS) domain

```
CC of PKAG3 and is useful in gene therapy.
XX
SQ Sequence 464 AA;

Query Match 94.5%; Score 2398; DB 22; Length 464;
Best Local Similarity 99.8%; Pred. No. 2.2e-239;
Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 26 MSFLEQENSSWSPAVTSSSRTGKRRAKALRWTRQKSVBEGPPGQGGPRSRPAAE 85
Db 1 MSFLEQENSSWSPAVTSSSRTGKRRAKALRWTRQKSVBEGPPGQGGPRSRPAE 60

Qy 86 STGLEATPPKTTPLAQADPAGVGTPTGWDCLPSDCTASAGSDDVELATEFPATEAW 145
Db 61 STGLEATPPKTTPLAQADPAGVGTPTGWDCLPSDCTASAGSDDVELATEFPATEAW 120

Qy 146 ECELEGLLEERPALCLSPQAPFKLGWDELKPKGAQIYMRFMQHTCYDAMATSSKLI 205
Db 121 ECELEGLLEERPALCLSPQAPFKLGWDELKPKGAQIYMRFMQHTCYDAMATSSKLI 180

Qy 206 FDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEIE 265
Db 181 FDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEIE 240

Qy 266 OHKIETWREIYLOGCKPLVSTSPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVHLIHTK 325
Db 241 OHKIETWREIYLOGCKPLVSTSPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVHLIHTK 300

Qy 326 RLLAFHLIFGSLLPSPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSLPV 385
Db 301 RLLAFHLIFGSLLPSPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSLPV 360

Qy 386 VNECGQVVGLYSRFDVHLLAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVID 445
Db 361 VNECGQVVGLYSRFDVHLLAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVID 420

Qy 446 RIAREQVHRLVVDVETQHLLGVVSLDILQALVLSPAGIDALGA 489
Db 421 RIAREQVHRLVVDVETQHLLGVVSLDILQALVLSPAGIDALGA 464

RESULT 3
AAE00224
ID AAE00224 standard; Protein; 514 AA.
XX
AC AAE00224;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa Pkag3 splice variant.
XX
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy.
XX
OS Sus scrofa.
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
```

```
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX WPI: 2001-244810/25.
DR N-PSDB; AAD03321.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 5; Page 70-71; 71pp; English.
XX
CC The present sequence is pig adenosine monophosphate (AMP)-activated
CC kinase (AMPK) gamma subunit muscle-specific isoform, PKAG3 splice
CC variant. Pkag3 gene is located in the RN locus of chromosome 15.
CC Mutation in Pkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PKAG3.
CC Transgenic animal and host cell transformed with PKAG3 or a
CC heterotrimeric AMPK consisting of PKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PKAG3 is useful for detecting mutations in a Pkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PKAG3 and is useful in gene therapy.
XX
SQ Sequence 514 AA;

Query Match 82.2%; Score 2085; DB 22; Length 514;
Best Local Similarity 82.0%; Pred. No. 7.5e-207;
Matches 418; Conservative 20; Mismatches 46; Indels 26; Gaps 3;

Qy 5 LEHALRR-----TPSSSLGGSEHOEHSFLEQENSSWPSP 40
Db 6 LEQALRRVPSRGNGWELEQLRPEGRGPTADTPSSSLGGPKHQSFSFLEQESKSPSR 65

Qy 41 AVTSSSERIRGKRRAKALRWTRQKSVBEGPPGQGGPRSRPAESTGLEATFPKTTPLA 100
Db 66 AVTSSSERHGDQGNKASRWTRQEDVEBGGPPGREGPQSRPVAESTGQEAFFPKATPLA 125

Qy 101 QADP-AGVGTPTGWDCLPSDCTASAGSDDVELATEFPATEAWCELEGLLEERPAL 159
Db 126 QAAPLAEDVNDPPTERDILPSCDAASASDSNTDHLDLGTEFSASAAGDEL-GLVEKPAP 184

Qy 160 CLSPQAPPKLGWDELKPKGAQIYMRFMQHTCYDAMATSSKLIFFDTMLEIKKAPAL 219
Db 185 CFSPEVLLPLUGWDELQPKGAQVYMHFMQHTCYDAMATSSKLIFFDTMLEIKKAPAL 244

Qy 220 VANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEIEBQHKIETWREIYLOG 279
Db 245 VANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEIEBQHKIETWREIYLOG 304

Qy 280 CFKPLVSTSPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVHLIHTKRLKLFHIFGSLLP 339
Db 305 CFKPLVSTSPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVHLIHTKRLKLFHIFGSLLP 364

Qy 340 RPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSLPVWNECGQVGLYSRF 399
Db 365 RPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSLPVWNECGQVGLYSRF 424

Qy 400 DVHILAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVIDRIARQVHRLVLD 459
Db 425 DVHILAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVIDRIARQVHRLVLD 484

Qy 460 ETQHLLGVVSLDILQALVLSPAGIDALGA 489
```

|||||  
485 ETQHLLGVVSLSDIIQLALVLSGIDALGA 514

## RESULT 4

AAE22985  
ID AAE22985 standard; Protein: 464 AA.

AC AAE22985;

DT 09-AUG-2002 (first entry)

XX pig PRKAG3 polymorphic variant (PRKAG3-30).

KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
KW variant.

XX Sus scrofa.

XX Key Location/Qualifiers

FT Misc-difference 30

FT /note= "Wild type Asn is substituted with Thr due  
to single nucleotide polymorphism (SNP)"

XX WO200220850-A2.

PN 14-MAR-2002.

XX 10-SEP-2001; 2001WO-US28283.

XX 08-SEP-2000; 2000US-231045P.

PR 08-JAN-2001; 2001US-260239P.

PR 18-JUN-2001; 2001US-299111P.

XX (IOWA ) UNIV IOWA STATE RES FOUND INC.

XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;

XX WPI; 2002-393850/42.

DR N-PSDB; AAD36457.

XX Screening animals to determine those likely to produce larger litters

PT and improved meat quality traits involves assaying for the presence of

PT polymorphisms in the AMP activated protein kinase regulatory gamma

PT subunit gene

XX Disclosure; Page 91-93; 109pp; English.

XX The invention relates to a method for screening animals to determine

CC those more likely to produce large litters and improved meat quality

CC traits. The method involves assaying for the presence of a genotype

CC in the sample of genetic material obtained from animal. The genotype

CC is characterised by polymorphism(s) in the AMP activated protein

CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used

CC for screening animals e.g., pigs to determine those most likely to

CC exhibit improved meat quality traits and to produce larger litters.

CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-30).

XX SQ Sequence 464 AA;

Query Match

Best Local Similarity 79.3%; Score 2013; DB 23; Length 464;

Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;

Qy 26 MSFLEQBNSSWSPSPAVTSSSRINGKRAKALRWTRQKSVBEGPQGGPRSPRAE 85

Db 1 MSFLEQGESRSPSRVATTTSSRSRSHGQGTAKSRWRQEDVEEGPPGPRGQSRPVAE 60

Qy 86 STGLEATPKTTPLAQADP-AGVGTPTTCHDCLPSCDTASAGSSTDDVELATEFPATEA 144

Db 61 STGQBATFPKATPLAQAAPLAEVDNPPPTERDILLPSDCAASADSNTDHLDLGIEFSASA 120

Qy 145 WECELEGLLEERPALCLSPQAPPKLGWMDDELKRPKGAQIYRMFMQEHCTCYDAMATSKLV 204  
Db 121 SGDEL-GLVEEKPAFCPSPEVLLPRLGWDDELQKPGAQVYMHFMQEHCTCYDAMATSKLV 179  
Qy 205 IFDTMLEIKKAPFALVANGVRAAPLWDSKQSFVGMLTITDFFILVLRHYRSPVQIYEI 264  
Db 180 IFDTMLEIKKAPFALVANGVRAAPLWDSKQSFVGMLTITDFFILVLRHYRSPVQIYEI 239  
Qy 265 EQHKIETWREIYLQGCCKPLVSI SPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTH 324  
Db 240 EEHKIETWREIYLQGCCKPLVSI SPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTH 299  
Qy 325 KRLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRLAVVLETPILTALDIFVDRVSALP 384  
Db 300 KRLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRLAVVLETPILTALDIFVDRVSALP 359  
Qy 385 VYNEGQVVGLYSRFDVIHLAAQOQYTNHLDMSVGEALRQRTLCLEGLVSCOPHESLGEVI 444  
Db 360 VYNETGVVGLYSRFDVIHLAAQOQYTNHLDMSVGEALRQRTLCLEGLVSCOPHETLGEVI 419  
Qy 445 DRIAREQVHRLVLVDQTHLLGVVSLSDIQLALVLSGIDALGA 489  
Db 420 DRIAREQVHRLVLVDQTHLLGVVSLSDIQLALVLSGIDALGA 464

## RESULT 5

AAE00222

ID AAE00222 standard; Protein: 464 AA.

AC AAE00222;

DT 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW Cystathione beta synthase; CBS; cardiac; gene therapy; RN locus;  
KW chromosome 15.

XX Sus scrofa.

EH Key Location/Qualifiers

FT Domain 172..225

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Misc-difference 200

FT /note= "RN- mutation site"

FT Domain 253..307

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Domain 329..382

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Domain 400..453

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

XX WO200120003-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP098996.

XX 10-SEP-1999; 99EP-0402236.

PR 18-MAY-2000; 2000EP-0401388.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA (ANDE/) ANDERSSON L.

PA (LOOF/) LOOFT C.

PA (KALM/) KALM E.

XX



PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
DR WPI: 2001-244810/25.  
DR N-PSDB; AAD03319.  
XX  
PT New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy .  
XX  
XX Claim 5: Fig 3: 7lpp; English.  
XX  
CC The present sequence is pig adenosine monophosphate  
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome  
CC 15. Mutation in Prkag3 results in an altered regulation of carbohydrate  
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
CC therapeutic for treating carbohydrate metabolism disorders such as  
CC diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
XX  
XX Sequence 464 AA;  
SQ  
Query Match 79.2%; Score 2011; DB 22; Length 464;  
Best Local Similarity 86.0%; Pred. No. 3e-199;  
Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;  
QY 26 MSFLEQENSSWSPSPAVTSSSRIRGKRKALRWTRQKSVSEGGPPQGGGPRSPRAE 85  
Db 1 MSFLEQESRSWSPAVTSSSRIRGKRKALRWTRQKSVSEGGPPQGGGPRSPRAE 60  
QY 86 STGLEATFPKTTPLAADI-ACVGTPTGMDCLPSPCTASAGSSTDDVELATEFPATEA 144  
Db 61 STGEATFPKATPLAQAAPLAELAVINPPTKDIPLSPCAASDSNTDHLGIEFFSAA 120  
QY 145 WECELEGLLEERFALCISQAPFPFKLGWDDLRKPAQIYMRFMQHTCYDAMATSSKLV 204  
Db 121 SGDEL-GLVEKFAPCSPSEVLLPRLGWDDLRKPAQIYMRFMQHTCYDAMATSSKLV 179  
QY 205 IFDTMLBKAFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVLQIYEI 264  
Db 180 IFDTMLBKAFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVLQIYEI 239  
QY 265 EOHKIEWRIYLOGCFKPLYSISPNDLSFEAVYTIKRIHRLVPLDPSGNNVLHLTH 324  
Db 240 EEHKIEWRIYLOGCFKPLYSISPNDLSFEAVYTIKRIHRLVPLDPSGNNVLHLTH 299  
QY 325 KRLKLFHIFGSLPRPSFLYRTIQDLGIGTFRDLAWLETAPILTALDIFVDRVSALP 384  
Db 300 KRLKLFHIFGTLPRPSFLYRTIQDLGIGTFRDLAWLETAPILTALDIFVDRVSALP 359  
QY 385 VYNECGQVGLYSRFDVHIAAQYTYNHLNMSVGEALRQTLCLSGVLSCOPHESLGEVI 444  
Db 360 VYNETGVGLYSRFDVHIAAQYTYNHLNMSVGEALRQTLCLSGVLSCOPHETLGEVI 419  
QY 445 DRIAREQVHRLVLVDQTHLLGVNLSLIDLOALVLSFAGIDALGA 489  
Db 420 DRIAREQVHRLVLVDQTHLLGVNLSLIDLOALVLSFAGIDALGA 464

RESULT 6  
AAE22984  
ID AAE22984 standard; Protein; 464 AA.  
XX  
AC AAE22984;  
XX  
DT 09-AUG-2002 (first entry)  
XX  
DE Pig wild-type PRKAG3 protein.  
XX  
KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
KW screening; meat quality; single nucleotide polymorphism; SNP; pig.  
XX  
OS Sus scrofa.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 30  
FT /note= "Wild type Asn is replaced with Thr during  
FT single nucleotide polymorphism (SNP)."  
FT Misc-difference 52  
FT /note= "Wild type Gly is replaced with Ser during  
FT single nucleotide polymorphism (SNP)."  
FT Misc-difference 199  
FT /note= "Wild type Val is replaced with Ile during  
FT single nucleotide polymorphism (SNP)."  
FT Misc-difference 200  
FT /note= "Wild type Arg is replaced with Gln during  
FT single nucleotide polymorphism (SNP)."  
XX  
XX WO200220850-A2.  
XX  
XX PN  
XX PD 14-MAR-2002.  
XX  
XX PF 10-SEP-2001; 2001WO-US28283.  
XX  
XX PR 08-SEP-2000; 2000US-231045P.  
XX PR 08-JAN-2001; 2001US-260239P.  
XX PR 18-JUN-2001; 2001US-299111P.  
XX  
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.  
XX  
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;  
XX  
XX WPI: 2002-393850/42.  
XX N-PSDB; AAD36456.  
XX  
XX Screening animals to determine those likely to produce larger litters  
XX and improved meat quality traits involves assaying for the presence of  
XX polymorphisms in the AMP activated protein kinase regulatory gamma  
XX subunit gene .  
XX  
XX Claim 2: Fig 1: 109pp; English.  
XX  
XX The invention relates to a method for screening animals to determine  
XX those more likely to produce large litters and improved meat quality  
XX traits. The method involves assaying for the presence of a genotype  
XX in the sample of genetic material obtained from animal. The genotype  
XX is characterised by polymorphism(s) in the AMP activated protein  
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
XX for screening animals e.g., pigs to determine those most likely to  
XX exhibit improved meat quality traits and to produce larger litters.  
XX The present sequence is pig wild-type PRKAG3 protein.  
XX  
SQ Sequence 464 AA;  
Query Match 79.2%; Score 2011; DB 23; Length 464;  
Best Local Similarity 86.0%; Pred. No. 3e-199;  
Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;  
QY 26 MSFLEQENSSWSPSPAVTSSSRIRGKRKALRWTRQKSVSEGGPPQGGGPRSPRAE 85  
Db 1 MSFLEQESRSWSPSPAVTSSSRIRGKRKALRWTRQKSVSEGGPPQGGGPRSPRAE 60



XX PD 14-MAR-2002.  
XX PF 10-SEP-2001; 2001WO-US28283.  
XX PR 08-SEP-2000; 2000US-231045P.  
PR 08-JAN-2001; 2001US-260239P.  
PR 18-JUN-2001; 2001US-299111P.  
XX PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
XX PI Rothschild MF, Ciobanu DC, Malek M, Piastow G;  
XX DR WPI: 2002-393850/42.  
DR N-PSDB; AAD36460.  
XX PT Screening animals to determine those likely to produce larger litters  
PT and improved meat quality traits involves assaying for the presence of  
PT polymorphisms in the AMP activated protein kinase regulatory gamma  
PT subunit gene  
XX PS Disclosure; Page 105-107; 109pp; English.  
XX CC The invention relates to a method for screening animals to determine  
CC those more likely to produce large litters and improved meat quality  
CC traits. The method involves assaying for the presence of a genotype  
CC in the sample of genetic material obtained from animal. The genotype  
CC is characterised by polymorphism(s) in the AMP activated protein  
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
CC for screening animals e.g., pigs to determine those most likely to  
CC exhibit improved meat quality traits and to produce larger litters.  
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).  
XX SQ Sequence 464 AA;

Query Match 79.1%; Score 2007; DB 23; Length 464;  
Best Local Similarity 85.8%; Pred. No. 7.8e-199;  
Matches 399; Conservative 20; Mismatches 44; Indels 2; Gaps 2;

QY 26 MSFLEQSSSWPSPAVTSSSRIRGKRAKALRWTRQKSVSEEGEPGQSGRSPRAE 85  
DB 1 MSFLEQSGRSWPSRAVTTSSSRSHGQGNKASRWTRQEDVEEGPGPGPSRPAE 60

QY 86 STGLEATFPKTTPLAQADP-AGVGTPTGWDCLPDSCTASAGSSTDVLEATFPATEA 144  
DB 61 STQGEATFPKATPLAQAPLAEDVNPPTERDILPSCDCAASASDNTDHLDLGIEFSASA 120

QY 145 WECELEGLLEERPALCLSPQAPPKPKLGDDELKPKGAQIYMRFMWQHTCYDAMATSKLV 204  
DB 121 SGDEL-GLVEEKPAFCPSFEVLLPRLGDDELQKGAQYVHFQWHTCYDAMATSKLV 179

QY 205 IFDTMLEIKKAFKAFVALVANGVRAAPLWDSKKQSFVGMLTITDIFILVHRYRSPVQIYEI 264  
DB 180 IFDTMLEIKKAFKAFVALVANGVRAAPLWDSKKQSFVGMLTITDIFILVHRYRSPVQIYEI 239

QY 265 EQHKIETWREIYLOCCFPLVSPNDLSFEAVYTLIKNRIHRLPVLDPVSGNVLHILPH 324  
DB 240 EEHKIETWREIYLOCCFPLVSPNDLSFEAVYTLIKNRIHRLPVLDPVSGAVLHILPH 299

QY 325 KRLLKELHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLETPAPILTALDIFVDRVSALP 384  
DB 300 KRLLKELHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLETPAPILTALDIFVDRVSALP 359

QY 385 VVNECGQVGLYSRFDVTHLAAQOYTNHLDMSVGEALRQRTLCLEGLVSCOPHESLGEVI 444  
DB 360 VVNETQVGLYSRFDVTHLAAQOYTNHLDMSVGEALRQRTLCLEGLVSCOPHETLGEVI 419

QY 445 DRIAREQVHRLVLDVETQHLGLGVVSLSDILQALVLSFAGIDALGA 489  
DB 420 DRIAREQVHRLVLDVETQHLGLGVVSLSDILQALVLSFAGIDALGA 464

RESULT 9

AAE22986  
ID AAE22986 standard; Protein; 464 AA.  
XX AC AAE22986;  
XX DT 09-AUG-2002 (first entry)  
XX DE Pig PRKAG3 polymorphic variant (PRKAG3-52).  
XX KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
KW variant.  
XX OS Sus scrofa.  
XX FH Key Location/Qualifiers  
FT Misc-difference 52  
FT /note= "Wild type Gly is substituted with Ser due  
FT to single nucleotide polymorphism (SNP)"  
XX PN WO2002220850-A2.  
XX PD 14-MAR-2002.  
XX PF 10-SEP-2001; 2001WO-US28283.  
XX PR 08-SEP-2000; 2000US-231045P.  
PR 08-JAN-2001; 2001US-260239P.  
PR 18-JUN-2001; 2001US-299111P.  
XX PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
XX PI Rothschild MF, Ciobanu DC, Malek M, Piastow G;  
XX DR WPI: 2002-393850/42.  
DR N-PSDB; AAD36458.  
XX PT Screening animals to determine those likely to produce larger litters  
PT and improved meat quality traits involves assaying for the presence of  
PT polymorphisms in the AMP activated protein kinase regulatory gamma  
PT subunit gene  
XX PS Claim 36; Page 96-97; 109pp; English.  
XX CC The invention relates to a method for screening animals to determine  
CC those more likely to produce large litters and improved meat quality  
CC traits. The method involves assaying for the presence of a genotype  
CC in the sample of genetic material obtained from animal. The genotype  
CC is characterised by polymorphism(s) in the AMP activated protein  
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
CC for screening animals e.g., pigs to determine those most likely to  
CC exhibit improved meat quality traits and to produce larger litters.  
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).  
XX SQ Sequence 464 AA;

Query Match 79.0%; Score 2005; DB 23; Length 464;  
Best Local Similarity 85.8%; Pred. No. 1.3e-198;  
Matches 399; Conservative 19; Mismatches 45; Indels 2; Gaps 2;

QY 26 MSFLEQSSSWPSPAVTSSSRIRGKRAKALRWTRQKSVSEEGEPGQSGRSPRAE 85  
DB 1 MSFLEQSGRSWPSRAVTTSSSRSHGQGNKASRWTRQEDVEEGPGPGPSRPAE 60

QY 86 STGLEATFPKTTPLAQADP-AGVGTPTGWDCLPDSCTASAGSSTDVLEATFPATEA 144  
DB 61 STQGEATFPKATPLAQAPLAEDVNPPTERDILPSCDCAASASDNTDHLDLGIEFSASA 120

QY 145 WECELEGLLEERPALCLSPQAPPKPKLGDDELKPKGAQIYMRFMWQHTCYDAMATSKLV 204  
DB 121 SGDEL-GLVEEKPAFCPSFEVLLPRLGDDELQKGAQYVHFQWHTCYDAMATSKLV 179

QY 205 IFDTMLEIKKAFKAFVALVANGVRAAPLWDSKKQSFVGMLTITDIFILVHRYRSPVQIYEI 264

```

Db 180 IFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEI 239
QY 265 EQHKIETWREIYLOGCEKPLVSI:SPNDLSFEAVYTLIKNRTHRLPVLDPVSGNVLHILTH 324
Db 240 EEHKIETWREIYLOGCEKPLVSI:SPNDLSFEAVYTLIKNRHRLPVLDPVSGAVLHILTH 299
QY 325 KRLKFLHIFGSLPLRPSEFLYRTIQDLGIGTFRDLAVVLETPAPILTALDIFVDRRVSA 384
Db 300 KRLKFLHIFGTLPLRPSEFLYRTIQDLGIGTFRDLAVVLETPAPILTALDIFVDRRVSA 359
QY 385 VVNECGQVGLYSRFDVIHLAAQOTYNHLDMSVGEALPQRTLCLEGLVSCOPHESLGEVI 444
Db 360 VVNETGQVWGLYSRFDVIHLAAQOTYNHLDMSVGEALPQRTLCLEGLVSCOPHETLGEVI 419
QY 445 DRIAREQVHRLVLDVDETHLGGVSLSDILQALVLSPAGIDALGA 489
Db 420 DRIAREQVHRLVLDVDETHLGGVSLSDILQALVLSPAGIDALGA 464

RESULT 10
AAE00221
ID AAE00221 standard; Protein: 305 AA.
XX
AC AAE00221;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human AMPK gamma subunit muscle-specific isoform, PRKAG3.
XX
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiac; gene therapy.
OS Homo sapiens.
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI: 2001-244810/25.
DR N-PSDB; AAE03296.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 4; Page 55-57; 71pp; English.
XX
CC The present sequence is human adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC useful as therapeutic for treating carbohydrate metabolism disorders such
CC as diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic

```

```

CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 305 AA;

```

Query Match 61.1%; Score 1551; DB 22; Length 305;

Best Local Similarity 100.0%; Pred. No. 7.7e-152;

Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 MRPMQEHCTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 244

Db 1 MRPMQEHCTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 60

QY 245 DFILVLRHYRSPVQIYETLQHKIETWREIYLOGCEKPLVSI:SPNDLSFEAVYTLIKNR 304

Db 61 DFILVLRHYRSPVQIYETLQHKIETWREIYLOGCEKPLVSI:SPNDLSFEAVYTLIKNR 120

QY 305 IHRPLVDPVSGNVLHILTHKRLKFLHIFGSLPLRPSEFLYRTIQDLGIGTFRDLAVVLE 364

Db 121 IHRPLVDPVSGNVLHILTHKRLKFLHIFGSLPLRPSEFLYRTIQDLGIGTFRDLAVVLE 180

QY 365 TAPILTALDIFVDRRVSAIPVNECGQVGLYSRFDVIHLAAQOTYNHLDMSVGEALPQR 424

Db 181 TAPILTALDIFVDRRVSAIPVNECGQVGLYSRFDVIHLAAQOTYNHLDMSVGEALPQR 240

QY 425 TLCLEGLVSCOPHESLGEVIDRIAREQVHRLVLDVDETHLGGVSLSDILQALVLS 484

Db 241 TLCLEGLVSCOPHESLGEVIDRIAREQVHRLVLDVDETHLGGVSLSDILQALVLS 300

QY 485 DALGA 489

Db 301 DALGA 305

RESULT 11

AAE00329

ID AAE00329 standard; Protein: 305 AA.

XX

AC AAE00329;

XX

DT 13-JUN-2001 (first entry)

XX

DE Human Prkag3 V40I mutant.

XX

KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;

KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;

KW cystathione beta synthase; CBS; cardiac; gene therapy; mutant; muten;

KW variant.

OS Homo sapiens.

XX

XX

FH Key Location/Qualifiers

FT Misc-difference 40

FT /note= "Wild-type Val substituted with Ile"

PN WO200120003-A2.

XX

PD 22-MAR-2001.

XX

PF 11-SEP-2000; 2000WO-EP09896.

XX

PR 10-SEP-1999; 99EP-0402236.

PR

PR 18-MAY-2000; 2000EP-0401388.



CC is derived from the human Prka3 sequence SEQ.ID.NO.4 shown in page 57-58  
XX of sequence listing (AAE00221).

SQ Sequence 305 AA;

Query Match 61.0%; Score 1547; DB 22; Length 305;  
Best Local Similarity 99.7%; Pred. No. 2e-151;  
Matches 304; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 185 MRFQEHCTCYDAMATSSKLVIFDWTMLEIKKAPFALVANGVRAAPLWDSKKSQSFVGMLTIT 244  
DB 1 MRFQEHCTCYDAMATSSKLVIFDWTMLEIKKAPFALVANGVRAAPLWDSKKSQSFVGMLTIT 60

QY 245 DFILVLRHYRSPVGVNHLTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 304

DB 61 DFILVLRHYRSPVGVNHLTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 120

QY 305 IHRPLVLDPVSGNVLHILTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 364

DB 121 IHRPLVLDPVSGNVLHILTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 180

QY 365 TAPILTALDIFVDRRYSALPVVNECGQVGLYSRFDVHILAAQQTYNHLDMSVGEALRQR 424

DB 181 TAPILTALDIFVDRRYSALPVVNECGQVGLYSRFDVHILAAQQTYNHLDMSVGEALRQR 240

QY 425 TLCLEGVLSQCPHESIGEVIDRIAREQVHRVLYVDETHQLLGVWSLSILQALVLSPAGI 484

DB 241 TLCLEGVLSQCPHESIGEVIDRIAREQVHRVLYVDETHQLLGVWSLSILQALVLSPAGI 300

QY 485 DALGA 489

DB 301 DALGA 305

RESULT 13

AAE00220

ID AAE00220 standard; Protein; 305 AA.

AC AAE00220;

DT 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3.

DE Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;

KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;

KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;

KW chromosome 15.

OS Sus scrofa.

Key Location/Qualifiers

Domain 13..66

/label= CBS

/note= "Cystathione beta synthase domain"

Domain 94..148

/label= CBS

/note= "Cystathione beta synthase domain"

Domain 170..223

/label= CBS

/note= "Cystathione beta synthase domain"

Domain 241..294

/label= CBS

WO200120003-A2.

22-MAR-2001.

11-SEP-2000; 2000WO-EF09896.

10-SEP-1999; 99EP-0402236.

18-MAY-2000; 2000EP-0401388.

(INRG ) INRA INST NAT RECH AGRONOMIQUE.

(ANDE/) ANDERSSON L.

(LOOF/) LOOFT C.

(KALM/) KALM E.

Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

Iannuccelli N, Gellin J, Le Roy P, Chardon P;

WPI: 2001-244810/25.

N-PSDB; AAD03295.

New variants of the gamma subunit of vertebrate adenosine

monophosphate-activated kinase for diagnosis or treatment of disorders

associated with energy metabolism such as diabetes, obesity, and

myopathy .

Claim 4: Fig 2: 7lpp: English.

The present amino acid sequence is pig adenosine monophosphate

(AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,

PRKAG3. Prka3 gene is located in the RN locus of chromosome 15.

Mutation in Prka3 results in an altered regulation of carbohydrate

metabolism, particularly in skeletal muscle. PRKAG3 is useful as

therapeutic for treating carbohydrate metabolism disorders such as

diabetes, obesity, and disorders associated with muscle metabolism

such as myopathy and cardiovascular diseases, to modulate AMPK

activity, and for restoring a normal AMPK function. PRKAG3 sequence

and its functionally altered mutants are useful for the diagnostic

evaluation, genetic testing and prognosis of a metabolic disorder,

preferably a carbohydrate metabolism disorder. Primers that can detect

a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

useful for detecting a dysfunction of carbohydrate metabolism resulting

from the expression of a functionally altered allele of PRKAG3.

Transgenic animal and host cell transformed with PRKAG3 or a

heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

screening compounds able to modulate AMPK activity. Nucleic acid

encoding PRKAG3 is useful for detecting mutations in a Prka3 gene, or

in a sequence encoding the first cystathione beta synthase (CBS) domain

of PRKAG3 and is useful in gene therapy.

SQ Sequence 305 AA;

Query Match 59.4%; Score 1507; DB 22; Length 305;

Best Local Similarity 97.0%; Pred. No. 2.8e-147;

Matches 296; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 185 MRFQEHCTCYDAMATSSKLVIFDWTMLEIKKAPFALVANGVRAAPLWDSKKSQSFVGMLTIT 244

DB 1 MRFQEHCTCYDAMATSSKLVIFDWTMLEIKKAPFALVANGVRAAPLWDSKKSQSFVGMLTIT 60

245 DFILVLRHYRSPVGVNHLTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 304

DB 61 DFILVLRHYRSPVGVNHLTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 120

QY 305 IHRPLVLDPVSGNVLHILTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 364

DB 121 IHRPLVLDPVSGNVLHILTHKLLKFLHIFGSLLPSPSYRTIQDLGIGTFRDLAVVLE 180

QY 365 TAPILTALDIFVDRRYSALPVVNECGQVGLYSRFDVHILAAQQTYNHLDMSVGEALRQR 424

DB 181 TAPILTALDIFVDRRYSALPVVNECGQVGLYSRFDVHILAAQQTYNHLDMSVGEALRQR 240

QY 425 TLCLEGVLSQCPHESIGEVIDRIAREQVHRVLYVDETHQLLGVWSLSILQALVLSPAGI 484

DB 241 TLCLEGVLSQCPHESIGEVIDRIAREQVHRVLYVDETHQLLGVWSLSILQALVLSPAGI 300

QY 485 DALGA 489

DB 301 DALGA 305

RESULT 14  
AAE00226  
ID AAE00226 standard; Protein; 305 AA.  
XX  
AC AAE00226;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).  
XX  
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT Domain 13..66  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Misc-difference 40  
FT /note= "Wild-type Val substituted with Ile"  
FT Domain 94..148  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 170..223  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 241..294  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
XX  
PN WO200120003-A2.  
XX  
PD 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-EP09896.  
XX  
PR 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX  
WPI; 2001-244810/25.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX  
PS Claim 9; Page -; 71pp; English.  
XX  
CC The present sequence is a V40I mutant of muscle-specific isoform of  
CC gamma subunit of adenosine monophosphate (AMP)-activated kinase  
CC (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in  
CC decreased glycogen content in pig skeletal muscle.  
CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
CC therapeutic for treating carbohydrate metabolism disorders such as  
CC diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
CC Note: The present sequence is not shown in the specification, but  
CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).  
XX  
SQ Sequence 305 AA;  
Query Match 59.3%; Score 1506; DB 22; Length 305;  
Best Local Similarity 96.7%; Pred No 3 6e-147;  
Matches 295; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 185 MRFMQHTCYDAMATSSKLIVFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 244  
Db 1 MHFMQHTCYDAMATSSKLIVFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 60  
QY 245 DFTLVHLRYRSPLVQIYEIEHQKIETWREIYLOGCFPLVISPNDLSFEAVYTLIKNR 304  
Db 61 DFTLVHLRYRSPLVQIYEIEEHKIETWREIYLOGCFPLVISPNDLSFEAVYTLIKNR 120  
QY 305 IHRPLVLDPVSGNVLHILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRLAVLVE 364  
Db 121 IHRPLVLDPVSGAVLHILTHKRLKFLHIFGTLPRPSFLYRTIQDLGIGTFRLAVLVE 180  
QY 365 TABILTALDIFVDRVSALPVVNECCQVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 424  
Db 181 TABILTALDIFVDRVSALPVVNETGVVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 240  
QY 425 TLCLGVLSQPHESLGEVDRIARQVHRLVLDVDETOHLLGVVSLDILQALVLSPAGI 484  
Db 241 TLCLGVLSQPHETLGEVDRIARQVHRLVLDVDETOHLLGVVSLDILQALVLSPAGI 300  
QY 485 DALGA 489  
Db 301 DALGA 305  
RESULT 15  
AAE00225  
ID AAE00225 standard; Protein; 305 AA.  
XX  
AC AAE00225;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).  
XX  
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT Domain 13..66  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Misc-difference 41  
FT /note= "Wild-type Arg substituted with Gln"  
FT Domain 94..148  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 170..223  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"

FT	Domain	241...294
FT	/label= CBS	
FT	/note= "Cystathione beta synthase domain"	
PX		
PN	WO200120003-A2.	
PD	22-MAR-2001.	
PX		
PF	11-SEP-2000; 2000WO-EF09896.	
PP	10-SEP-1999; 99EP-0402236.	
PR	18-MAY-2000; 2000EP-0401388.	
PX		
PX	(INRG ) INRA INST NAT RECH AGRONOMIQUE.	
PA	(ANDE/) ANDERSSON L.	
PA	(LOOF/) LOOFT C.	
PA	(KALM/) KALM E.	
PI	Andersson L, Loof C, Kalim E, Milan D, Robic A, Rogel-Gaillard C;	
PI	Iannuccelli N, Gellin J, Le Roy P, Chardon P;	
XX	WPI; 2001-244810/25.	
PX		
PT	New variants of the gamma subunit of vertebrate adenosine	
PT	monophosphate-activated kinase for diagnosis or treatment of disorders	
PT	associated with energy metabolism such as diabetes, obesity, and	
PT	myopathy -	
PX		
XX	Claim 9; Page -: 71pp; English.	
PX		
CC	The present sequence is a R41Q mutant of muscle-specific isoform of	
CC	gamma subunit of adenosine monophosphate (AMP)-activated kinase	
CC	(AMPK) Prkag3 from Sus scrofa. This mutant sequence results in	
CC	increased glycogen content in pig skeletal muscle.	
CC	Mutation in Prkag3 results in an altered regulation of carbohydrate	
CC	metabolism, particularly in skeletal muscle. PRKAG3 is useful as	
CC	therapeutic for treating carbohydrate metabolism disorders such as	
CC	diabetes, obesity, and disorders associated with muscle metabolism	
CC	such as myopathy and cardiovascular diseases, to modulate AMPK	
CC	activity, and for restoring a normal AMPK function. PRKAG3 sequence	
CC	and its functionally altered mutants are useful for the diagnostic	
CC	evaluation, genetic testing and prognosis of a metabolic disorder,	
CC	preferably a carbohydrate metabolism disorder. Primers that can detect	
CC	a genetic polymorphic marker linked to a sequence encoding PRKAG3, are	
CC	useful for detecting a dysfunction of carbohydrate metabolism resulting	
CC	from the expression of a functionally altered allele of PRKAG3.	
CC	Transgenic animal and host cell transformed with PRKAG3 or a	
CC	heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for	
CC	screening compounds able to modulate AMPK activity. Nucleic acid	
CC	encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or	
CC	in a sequence encoding the first cystathione beta synthase (CBS) domain	
CC	of PRKAG3 and is useful in gene therapy.	
CC	Note: The present sequence is not shown in the specification, but	
CC	is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).	
XX		
SQ	Sequence 305 AA;	
	Query Match 59.2%; Score 1503; DB 22; Length 305;	
	Best Local Similarity 96.7%; Pred. No. 7.3e-147;	
	Matches 295; Conservative 5; Mismatches 5; Indels 0; Gaps 0;	
QY	185 MRFOEHCTCDAMATSSKLIVDFMTLEIKAFALVANGVAAAPLWDSSKQSFGMLTIT 244   1 MHFMQEHCTCDAMATSSKLIVDFMTLEIKAFALVANGVAAAPLWDSSKQSFGMLTIT 60	
DB		
QY	245 DFILVLHRYRSPLVQIYEIEBQHKTETWRIEYLQQCFPLVSIPNDSLSEFYVTLTKNR 304   61 DFILVLHRYRSPLVQIYEIEEHKTETWRIEYLQQCFPLVSIPNDSLSEFYVTLTKNR 120	
DB		
QY	305 IHRLPVLDPVSGNVHLITHLKRLKFUHFISGLLRPSPFLYRTIQDLIGITFRDLAVLVE 364   121 IHRLPVLDPVSGNAVLIHTLHKRLKFUHFISGLLRPSPFLYRTIQDLIGITFRDLAVLVE 180	
DB		



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:58:29 ; Search time 31.6865 Seconds  
(without alignments)  
1483.588 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2538  
Sequence: 1 MEPCLEHALRRTPSNSSLGG.....LSDILQALVLSGADALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1028	40.5	330	2 T10759	AMP-activated prote
2	488	19.2	478	2 T25899	hypothetical prote
3	457	18.0	322	1 RGYC3	regulatory protein
4	375.5	14.8	274	2 S67444	probable 5'-AMP-ac
5	243.5	9.6	391	2 B86222	hypothetical prote
6	222	8.7	443	2 T25854	hypothetical prote
7	217.5	8.6	460	2 T24248	hypothetical prote
8	214	8.4	379	2 T10971	pv42 protein - kid
9	198	7.8	424	2 T46197	hypothetical prote
10	196	7.7	629	2 T18227	hypothetical prote
11	189.5	7.5	352	2 F86287	hypothetical prote
12	179	7.1	399	2 D96832	hypothetical prote
13	163	6.4	447	2 B96720	hypothetical prote
14	161.5	6.4	392	2 G75056	dehydrogenase PAB0
15	158	6.2	527	2 S46088	hypothetical prote
16	155.5	6.1	392	2 D71175	hypothetical prote
17	149	5.9	280	1 H64452	conserved hypotnet
18	148.5	5.9	250	2 C90449	conserved hypotnet
19	148	5.8	92	2 T50087	probable 5'-AMP-ac
20	146	5.8	313	1 E69030	conserved hypotnet
21	141.5	5.6	157	1 F69185	yhcV homolog MTH64
22	139	5.5	152	2 AF2282	hypothetical prote
23	139	5.5	300	1 S08244	conserved hypotnet
24	138.5	5.5	421	2 C64475	hypothetical prote
25	138	5.4	485	1 JC4998	IMP dehydrogenase
26	137	5.4	284	1 H69355	conserved hypotnet
27	136.5	5.4	527	2 S64060	probable membrane
28	134	5.3	1286	2 A12195	two-component hybr
29	130.5	5.1	792	2 A84308	chloride channel [

30	129.5	5.1	1344	2 AD2103	two-component hybr
31	126.5	5.0	168	1 A64478	hypothetical prote
32	126	5.0	486	1 E71456	IMP dehydrogenase
33	125.5	4.9	157	2 AF3344	IMP dehydrogenase
34	123.5	4.9	287	2 F72780	hypothetical prote
35	122	4.8	282	2 C75148	hypothetical prote
36	120.5	4.7	223	2 T36952	conserved hypotnet
37	120.5	4.7	278	2 A72781	probable acetoin u
38	120.5	4.7	284	1 H69232	probable acetoin u
39	119	4.7	122	2 B69119	yhcV homolog MTH18
40	119	4.7	485	2 E75015	IMP dehydrogenase
41	116.5	4.6	382	2 C97543	hypothetical prote
42	116.5	4.6	382	2 AD2762	conserved hypotnet
43	115.5	4.6	285	1 C71188	conserved hypotnet
44	115.5	4.6	895	2 T03446	probable transcrip
45	115	4.5	482	2 C72264	IMP dehydrogenase

ALIGNMENTS

RESULT 1

T10759

AMP-activated protein kinase (EC 2.7.1.-) gamma chain - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000

C:Accession: T10759

R:Woods, A.; Cheung P, C.F.; Smith, F.C.; Davidson, M.D.; Scott, J.; Beri, R.K.; Carl

J. Biol. Chem. 271, 10282-10290, 1996

A:Title: Characterization of AMP-activated protein kinase beta and gamma subunits: p

A:Reference number: 206738; MUID:96215327; PMID:8626596

A:Accession: T10759

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-330 <WOO>

A:Cross-references: EMBL:X95578; NID:g1185270; PIDN:CAA64831.1; PID:g1185271

A:Experimental source: strain Wistar

C:Complex: heterotrimer; alpha, beta and gamma chains

C:Function:

A:Description: is responsible for the regulation of fatty acid synthesis by phosphor

C:Superfamily: CAT3 protein

C:Keywords: fatty acid biosynthesis; phosphotransferase

Query Match 40.5%; Score 1028; DB 2; Length 330;

Best Local Similarity 62.3%; Pred. No. 4.1e-72;

Matches 197; Conservative 56; Mismatches 63; Indels 0; Gaps 0;

QY 165 APFPLGHDDDELKPGAGIYMRFMQEHCTCYDAMATSSKLVIFDTMLEIKKAFALVANGV 224

Db 9 APAPENHSQETPSNSSVYTFMKSRCYDLIPTSSKLVVFTSLQVKKAFALVTNGV 68

QY 225 RAAPLWDSKQSFVGMLTITDFTLVHRYRSPVQIVEIQHKTREIYLQCGPKPL 284

Db 69 RAAPLWDSKQSFVGMLTITDFTLVHRYRSPVQIVEIQHKTREIYLQCGPKPL 128

QY 285 VQISPNDSLFAVYTLTKNRIHRLPVLDPVSGNVHLHTLTKRLKFLHIFGSLPRPSFL 344

Db 129 VCISPNASLFDVAVSSLRINKIHRPLVIDPESGNTLYLTTHKRLKFLKLFITEFPKPEFM 188

QY 345 YFTIQDLGIGTFRDLAVVLETAIPILTALDIFVDRVRVALPVNCCGVVGLSRDVIHL 404

Db 189 SXSLEELQIGYANIAMVTRTFVYVVGIGFVQHRVSALPVWDEKGRVVDIYSKFDVINL 248

QY 405 AAQQTYNHLDMSVGEALRQRTLCLEGVLSQCPHSEGLVIDRIARQVHRLVLYVDEOHL 464

Db 249 AAEKTYNLDVSVTKALQHSYFEGVLCXYLHETLEAILNRLVVEAEVRLVVDDEHDV 308

QY 465 LGWVSLSDILQALVLS 480

Db 309 KGIVSLSDILQALVLT 324

RESULT 2

T25899  
hypothetical protein T20F7.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T25899  
R:Miller, N.; Gattlung, S.  
submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of C. elegans cosmid T20F7.  
A:Reference number: T20107  
A:Accession: T25899  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-478 <MIL>  
A:Cross-references: EMBL:U97550; PIDN:AA852856.1; GSPDB:GN00028; CESP:T20F7.6  
A:Experimental source: strain Bristol N2; clone T20F7  
C:Genetics:  
A:Gene: CESP:T20F7.6  
A:Map position: X  
A:Introns: 33/3; 112/2; 144/1; 205/3; 263/2; 300/3; 380/3; 402/2; 451/1

Query Match 19.2%; Score 488; DB 2; Length 478;  
Best Local Similarity 36.7%; Pred. No. 5.4e-30;  
Matches 119; Conservative 69; Mismatches 106; Indels 30; Gaps 8;

QY 182 QIYRFRQHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSK-KQSFVGM 240  
Db 38 EAPARLLWNOQCEAMPSSKMWVDFQGLLMHKAFAKLLAQSTRHVLLSDPDPFGCKLDGI 97

QY 241 LTITDFILVLRHYRS-----PLVQI--YEIEQHKIETWRE-IYLGCGFKPLVS 286  
Db 98 LSVTDFIKVMLKYYRETRKCEKESFELDMTQIANEEIGNLSIRQYRELVKKEGNRLPLVS 157

QY 287 ISPNDLSFEAVYFLIKNRHRLPVLDVSGNVLLHILTHRLKLFHIFGSLLRPPSPLYR 346  
Db 158 VDASGSLGDAACILAEHVRHPIVDPDGSALFILTHKRLKFLWLFEGKFLAPLEYLHK 217

QY 347 TIQDLGTDFRLAVVLETAPILTALDIFVDRRVSAALPVV-NECGVGLYSRFDVHLA 405  
Db 218 SPKELGIGTWSGRVFPPTQVLCDDLIILKNKGVSLPVVERETFKVDMTISRFDAVGIA 277

QY 406 AQOYTNHLDMSVGEALRQRTLCLEG-----VLSCQPHESLGEVIDRIAREQVHRLVL 457  
Db 278 LE---NRLLDITVKEALAFKS---QGGPKMNDERVSVVRDNESEFKAVNVLDVHNVRLCA 331

QY 458 VDETQHLLGVVSLDILQALVSP 481  
Db 332 VNEHGGIEGVISLSDVINPMVWQP 355

RESULT 3  
KGBYC3  
regulatory protein SNF4 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: CAT3 protein; protein G2945; protein YGL115W  
C:Species: Saccharomyces cerevisiae  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 21-Jul-2000  
C:Accession: A38906; J03316; S48508; S64125  
R:Celenza, J.L.; Eng, F.J.; Carlson, M.  
Mol. Cell. Biol. 9, 5045-5054, 1989  
A:Title: Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae: evidence for  
A:Reference number: A33480; MUID:90097921; PMID:2481228  
A:Accession: A38906  
A:Molecule type: DNA  
A:Residues: 1-322 <CEL>  
A:Cross-references: GB:M30470; NID:g172635; PIDN:AAA35061.1; PID:g172636  
R:Schueller, H.J.; Entian, K.D.  
Gene 67, 247-257, 1988  
A:Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucose  
A:Reference number: J03316; MUID:89006284; PMID:3049255  
A:Accession: J03316  
A:Molecule type: DNA  
A:Residues: 1-322 <SCH>  
A:Cross-references: GB:M21760; NID:g171164; PIDN:AAA34472.1; PID:g171165  
R:Doi, A.; Doi, K.

submitted to the EMBL Data Library, June 1993  
A:Description: Correct end of the ORF for the CDC20 gene of Saccharomyces cerevisiae  
A:Reference number: S48507  
A:Accession: S48508  
A:Molecule type: DNA  
A:Residues: 1-21 <DOI>  
A:Cross-references: EMBL:D16506; NID:g391938; PIDN:BAA03958.1; PID:g2160324  
R:Laquin, G.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64122  
A:Accession: S64125  
A:Molecule type: DNA  
A:Residues: 1-322 <LAU>  
A:Cross-references: EMBL:T27637; NID:g1322666; PIDN:CAA96823.1; PID:g1322667; GSPDB:  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:SNF4; CAT34; MIPS:YGL115W  
A:Cross-references: SGD:S0003083; MIPS:YGL115W  
A:Map position: 7L  
C:Function:  
A:Description: involved in derepression of glucose-repressed genes  
C:Superfamily: CAT3 protein  
C:Keywords: nucleus; transcription regulation

Query Match 18.0%; Score 457; DB 1; Length 322;  
Best Local Similarity 34.2%; Pred. No. 7.9e-28;  
Matches 106; Conservative 66; Mismatches 108; Indels 30; Gaps 6;

QY 186 RPMOHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKQSFVGMITTD 245  
Db 24 RFLNSKTSYDVLVPSYRLVLDTSLLVKKSLNVLQNSIVSAPLWDSKTSFAGLLITTD 83

QY 246 FTLVLRHYRSPVQIYEIEQHKIETWREIYLGCGFKPL-----VSI SPNDLSLF 294  
Db 84 FINVIQYFNSP-----DKFELVKQLQDG-LADIERALGVLDQDTSIHPSRPLF 133

QY 295 EAVYTLKNRHRLPVLDVPSGN---VLHVLTHRLKLLKHLHIFGSLLRPS-FLYRTIQ 349  
Db 134 EACLMEKLSRSGRIPLIDQDETHREIYVSVLTQYRLKFEV----ALNCRHFHLKIPIG 189

QY 350 DLIGITGTRDLAVVLETAPILTALDIFVDRRVSAALPVVNECGVGLYSRFDVHLAAQQT 409  
Db 190 DLNIIITQDNKMSQMTPTVIDVIQMLTQGRVSSVPIIDENGVLINVEAYDVLGKGGI 249

QY 410 YNHLDMSVGEALRQRTLCLEGSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVWS 469  
Db 250 YNDLSLSVGEALRRSRDDFEGVYTCCKNDKLSLTMDNIRKARVHRFFVVDVGLVGLT 309

QY 470 LSDILQALVL 479  
Db 310 LSDILKVIILL 319

RESULT 4  
S67444  
probable 5'-AMP-activated, gamma subunit family - fission yeast (Schizosaccharomyces  
C:Species: Schizosaccharomyces pombe  
C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000  
C:Accession: T38059; S67444  
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: T21766  
A:Accession: T38059  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <MC2>  
A:Cross-references: EMBL:T65944; NID:g1217974; PIDN:CAA93805.1; PID:g1217975; GSPDB:  
C:Genetics:  
A:Map position: 1  
A:Introns: 12/3; 54/1; 173/3; 226/1  
C:Superfamily: CAT3 protein

Query Match 14.8%; Score 375.5; DB 2; Length 274;





Query Match 7.5%; Score 189.5; DB 2; Length 352;  
Best Local Similarity 23.2%; Pred. No. 5.1e-07;  
Matches 81; Conservative 59; Mismatches 128; Indels 81; Gaps 13

QY 192 TCYDAMATSKLVIFDTMLEIKKFAFFALVANGRAAPL-----W-----DSKKOS 236  
| | : : || : : |||| : : | : : | : : | : : | : : | : : | : : | : :  
Db 17 TAKDLTVNRRLVEVPYYTATISHAMNTLVANSISALPVAAPPGHWIAGGSGMIMESDKOT 76  
-----FVGMLITFDIL-----VLHYRYSPLVQIYEIBQHKE-----TWREIYLQC 280  
: : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 77 GVVRKHVIGILTMDLAHIAHGDSNLSDDRKMSSQVSSIIHCLEGLSLW-----128  
QY 281 FKPLVISPNDSLFEAVYTLTKNIHR--LPVLDPVSGN-----VLHILTHK 325  
::: || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 129 -----TLNPPTISLSECMEEVFSKG-IHRALPVPESSIESNNTIAGVELIESASAYKMLTQM 182  
QY 326 RLKLFL--HFGSLLPRPSFYRTIODLGIGTFERDLAVVLETAPTLTALDIFDVDRVSAL 383  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 193 DLLRFLKDHHFDOL---KIVLSRSISDLG-AVNDSVIATERFTVSNAINVMKALLNAV 238  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
QY 384 PVWN-----ECGVVGLYSRFVDIHAAQQTNYNHLDMSVGEALRQRTLCL 428  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 239 PIVHAPDIADHLQLVNGRHUKVKIGTSATLDLKGCRLEPQLTQLPTALEFTKTSKG 298  
QY 429 EGYLSGOPHSELGEVIDRIAREQVHRLVIVDETOHLLGWLSLSILOAL 477  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 299 REVVSCGVSTMEEAIEKVVTRGVHVWMDOQLLQGVVSLTDIIRSL 347  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
  
RESULT 12  
D96832  
hypothetical protein F18B13.17 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D96832  
R:Theologos, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; U.  
ansen, N.F.; Hughes, B.; Huizlar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.;  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; NUID:21015719; PMID:11130712  
A:Accession: D96832  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-399 <SFO>  
A:Cross-references: GB:AEO05173; NID:G5902384; PIDN:AAD55486.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F18B13.17  
A:Map position: 1  
  
Query Match 7.1%; Score 179; DB 2; Length 399;  
Best Local Similarity 25.3%; Pred. No. 3.9e-06;  
Matches 96; Conservative 59; Mismatches 118; Indels 106; Gaps 21

QY 182 QIYMRFNQEHCTCYDAMATSSKLV-----IFD-----TMLEIKKAFALVANGVRAAPL 229  
| | : : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 37 QSYSRLMQ-FKYKDLMDKRDLRVLPONATFLGDALNTMTILGIKP-----VANRVRAEV 90  
-----W-----DSKKOSFGMLTTIDFIL-----VLHYRYSPLV 259  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 91 AAKEPQWLGAAGSMIVELDKOSSARKQYIGWTMLDVVAHIADGESGLDKMAAPVS 150  
QY 260 QIYEIOKHETWRETYLQQCFKLVSISPNDSLFEAVYTLTKNIHRPLPDPVSGNVL 319  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 151 SLIGHCPGEGSLW-----SLNPNTSIMDCHEMLS-KG-IHR--VLYPLDSNTE 194  
QY 320 HILTHKRLKFHFGSL--LPRESFYLRITODL-GI--GTFRDLAVVLETAFLT----- 370  
| : || : : | : : | : : | : : | : : | : : | : : | : : | : : | : :





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 17.9842 Seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2538  
Sequence: 1 MEPGLEALRRTPSWSLG.....LSDILQALVLSPGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2398	94.5	464	1	AAKL_HUMAN
2	2011	79.2	464	1	AAKL_PIG
3	1093	43.1	569	1	AAKL_HUMAN
4	1028	40.5	330	1	AAKG_RAT
5	1017	40.1	331	1	AAKG_HUMAN
6	1013	39.9	330	1	AAKG_BOVIN
7	997	39.3	330	1	AAKG_MOUSE
8	457	18.0	322	1	SNF4_YEAST
9	456.5	18.0	334	1	YDAL_SCHPO
10	418.5	16.5	328	1	SNF4_KLULA
11	267.5	10.5	133	1	AAKG_PIG
12	158	6.2	527	1	YB64_YEAST
13	149	5.9	280	1	YC25_METJA
14	139	5.5	300	1	YR33_THERPE
15	138.5	5.5	421	1	YE04_METJA
16	138	5.4	485	1	IMDH_PYRFU
17	136.5	5.4	527	1	YGF6_YEAST
18	126.5	5.0	168	1	Y26_METJA
19	126	5.0	486	1	IMDH_PYRHO
20	119	4.7	485	1	IMDH_PYRHO
21	114	4.5	1053	1	CAPP_SYN6
22	113.5	4.5	296	1	YC32_METJA
23	110	4.3	496	1	IMDH_METJA
24	108.5	4.3	944	1	VGLB_HSVT2
25	107.5	4.2	509	1	Y100_METJA
26	104.5	4.1	196	1	Y525_METKA
27	104	4.1	329	1	Y259_HELPY
28	103	4.1	537	1	IMDH_DROME
29	102.5	4.0	2377	1	CCAG_HUMAN
30	102	4.0	138	1	Y922_METJA
31	101.5	4.0	424	1	Y146_MYCGE
32	101.5	4.0	835	1	UL52_HSVSA
33	101	4.0	521	1	IMDH_CHLVI

## RESULT 1

ID	AAKL_HUMAN	STANDARD;	PRT;	464 AA.
AC	Q9UG19; Q9NRL1;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)			
DE	(AMPK gamma3).			
GN	PRKAG3 OR AMPKG3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20164049; PubMed=10698692;			
RA	Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;			
RT	Characterization of AMP-activated protein kinase gamma-subunit			
RT	isoforms and their role in AMP binding. #;			
RL	Biochem. J. 346:659-669(2000).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RP	TISSUE=Skeletal muscle;			
RX	MEDLINE=20280150; PubMed=10818001;			
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,			
RA	Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,			
RA	Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,			
RA	Andersson L.;			
RT	"A mutation in PRKAG3 associated with excess glycogen content in pig			
RT	skeletal muscle. #;			
RL	Science 288:1248-1251(2000).			
CC	!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID			
CC	SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO			
CC	REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND			
CC	INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-			
CC	SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE			
CC	IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.			
CC	!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A			
CC	GAMMA NON-CATALYTIC REGULATORY SUBUNITS.			
CC	!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART			
CC	AND PANCREAS.			
CC	!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA			
CC	SUBUNIT FAMILY.			
CC	!- SIMILARITY: CONTAINS 4 CBS DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AJ249977; CAB65117.1; ALT_INIT.			
DR	EMBL; AF214519; AAF73987.1; -.			



```
DR Genew: HGNC:9387; PRKAG3.
DR MM: 604976; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat: CBS domain.
FT DOMAIN 147 201 CBS 1.
FT DOMAIN 228 282 CBS 2.
FT DOMAIN 303 356 CBS 3.
FT DOMAIN 375 428 CBS 4.
FT CONFLICT 58 58 T -> A (IN REF. 1).
FT CONFLICT 163 164 MO -> IE (IN REF. 1).
FT CONFLICT 398 398 Q -> K (IN REF. 1).
FT CONFLICT 461 464 ALGA -> PSGPKI (IN REF. 1).
SQ SEQUENCE 464 AA; 51514 MW; 53985C2C77003A63 CRC64;

Query Match 94.5%; Score 2398; DB 1; Length 464;
Best Local Similarity 99.8%; Pred. No. 3.7e-174;
Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 MSFLEQENSSWSPSPAVTSSSERIRGKRAKALRWTRQKSVBEGPPQGQGGPRSPAAE 85
Db 1 MSFLEQENSSWSPSPAVTSSSERIRGKRAKALRWTRQKSVBEGPPQGQGGPRSPAAE 60

QY 86 STGLEATFPKTTPLAQADPAGVGTPTGWDCLPSDCTASAAGSSDDDELATEPPATEAW 145
Db 61 STGLEATFPKTTPLAQADPAGVGTPTGWDCLPSDCTASAAGSSDDDELATEPPATEAW 120

QY 146 ECEGLEGLEERPALCLSPQAPFPKLGWDELKRPQAGIYMRPMQEHCTCYDAMATSSKLVI 205
Db 121 ECEGLEGLEERPALCLSPQAPFPKLGWDELKRPQAGIYMRPMQEHCTCYDAMATSSKLVI 180

QY 206 FDMMLETKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVQIYEIE 265
Db 181 FDMMLETKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVQIYEIE 240

QY 266 QHKIETWRETYLQGCFFKPLVSIISPNDLSLFEAVVTLKNIHRLPVLDPVSGNVHLITHK 325
Db 241 QHKIETWRETYLQGCFFKPLVSIISPNDLSLFEAVVTLKNIHRLPVLDPVSGNVHLITHK 300

QY 326 RLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRVSALPV 385
Db 301 RLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRVSALPV 360

QY 386 VNECGQVGLYRSFQVHLAQQTYNHLDSVGEALNQRITLCLEGLSCQPHESLGEVID 445
Db 361 VNECGQVGLYRSFQVHLAQQTYNHLDSVGEALNQRITLCLEGLSCQPHESLGEVID 420

QY 446 RIAREQVHRLVLVDETHLGLGVSI SDILQALVLSPAGIDALGA 489
Db 421 RIAREQVHRLVLVDETHLGLGVSI SDILQALVLSPAGIDALGA 464

RESULT 2
AAKI.PIG
ID AAKI.PIG STANDARD; PRT; 464 AA.
AC Q9WT44.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)
DE (AMPK gamma3).
GN PRKAG3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT RN(-) GLN-200.
RC TISSUE=Skeletal muscle;
RA MEDLINE=20280150; PubMed=10818001;
RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,
RA Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,
```

```
RA Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,
RA Andersson L.;
RA *A mutation in PRKAG3 associated with excess glycogen content in pig
RT skeletal muscle.";
RL Science 288:1248-1251(2000).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE
CC IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -!- TISSUE SPECIFICITY: MUSCLE.
CC -!- DISEASE: DEFECTS IN PRKAG3 (RN-) ARE ASSOCIATED WITH EXCESS
CC GLYCOGEN CONTENT (ABOUT 70%) IN SKELETAL MUSCLE. THIS MUTATION
CC ORIGINATED IN THE HAMPSHIRE BREED PIGS AND HAS BENEFICIAL EFFECTS
CC ON MEAT CONTENT BUT DETRIMENTAL EFFECTS ON PROCESSING YIELD. THUS,
CC THIS MUTATION IS OF CONSIDERABLE ECONOMIC SIGNIFICANCE IN THE PIG
CC BREEDING INDUSTRY.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF214521; AAF73989.1; -.
DR EMBL: AF214520; AAF73988.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat: CBS domain; Disease mutation.
DR DOMAIN 172 226 CBS 1.
FT DOMAIN 253 307 CBS 2.
FT DOMAIN 328 381 CBS 3.
FT DOMAIN 400 453 CBS 4.
FT VARIANT 200 200 R -> Q (IN RN-).
SQ SEQUENCE 464 AA; 51308 MW; 17638CB12A2BA9DF CRC64;

Query Match 79.2%; Score 2011; DB 1; Length 464;
Best Local Similarity 86.0%; Pred. No. 7.4e-145;
Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;

QY 26 MSFLEQENSSWSPSPAVTSSSERIRGKRAKALRWTRQKSVBEGPPQGQGGPRSPAAE 85
Db 1 MSFLEQENSSWSPSPAVTSSSERIRGKRAKALRWTRQKSVBEGPPQGQGGPRSPAAE 60

QY 86 STGLEATFPKTTPLAQADP-AGVGTPTGWDCLPSDCTASAAGSSDDDELATEPPATEA 144
Db 61 STGLEATFPKTTPLAQADP-AGVGTPTGWDCLPSDCTASAAGSSDDDELATEPPATEA 120

QY 145 WECEGLEGLEERPALCLSPQAPFPKLGWDELKRPQAGIYMRPMQEHCTCYDAMATSSKLVI 204
Db 121 SGDEL-GLVEEKPAFCPSPEVLLPRLGWDELQKPGQAVYMHFMQEHCTCYDAMATSSKLVI 179

QY 205 IFDFTMLEIKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVQIYEIE 264
Db 180 IFDFTMLEIKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPVQIYEIE 239

QY 265 EOHKIETWRETYLQGCFFKPLVSIISPNDLSLFEAVVTLKNIHRLPVLDPVSGNVHLITHK 324
Db 240 EOHKIETWRETYLQGCFFKPLVSIISPNDLSLFEAVVTLKNIHRLPVLDPVSGNVHLITHK 299

QY 325 KRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRVSALPV 384
Db 300 KRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRVSALPV 359
```

QY 395 VNECGQVGLYSRQVDFVHLAAQQTYNHLDMSVGEALRQRTLCRLGVSLSCOPHESLGEVI 444  
 DB 360 VYNETGVGLYSRQVDFVHLAAQQTYNHLDMSVGEALRQRTLCRLGVSLSCOPHETLGEVI 419  
 QY 445 DRIAREQVRLVDETQHLGLGVSLSDILOALVLSFAGIDALGA 489  
 DB 420 DRIAREQVRLVDETQHLGLGVSLSDILOALVLSFAGIDALGA 464

RESULT 3  
 AAKH\_HUMAN  
 AC Q9UGJ0: Q9UGJ0; STANDARD; PRT; 569 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-2 subunit (AMPK gamma-2 chain)  
 DE (AMPK gamma2) (H91620p).  
 GN PRKAG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM A).  
 RX MEDLINE=20164049; PubMed=10698692;  
 RA Cheung P.C.F., Salt I.P., Davies S.P., Hardie D.G., Carling D.;  
 RT "Characterization of AMP-activated protein kinase gamma-subunit  
 RT isoforms and their role in AMP binding.";  
 RL Biochem. J. 346:659-669(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RX MEDLINE=20564210; PubMed=11112354;  
 RA Lang T., Yu L., Qiang T., Jiang J., Chen Z., Xin Y., Liu G., Zhao S.;  
 RT "Molecular cloning, genomic organization, and mapping of PRKAG2, a  
 RT heart abundant gamma-2 subunit of 5'-AMP-activated protein kinase, to  
 RT human chromosome 7q36.";  
 RL Genomics 70:258-263(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RC TISSUE=Placenta;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 218-569 FROM N.A.  
 RA Hattori A., Seki N., Hayashi A., Kozuma S., Muramatsu M., Saito T.;  
 RT "Human homolog of AMPK gamma-1 chain.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE OF 254-569 FROM N.A.  
 RA Walker C., Scott K., Bauer K., Harkins R.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP VARIANT WPWS GLY-531.  
 RX MEDLINE=21614537; PubMed=11748095;  
 RA Gollob M.H., Seger J.J., Gollob T.N., Tapscott T., Gonzales O.,  
 RA Bachinski L., Roberts R.;  
 RT "Novel PRKAG2 mutation responsible for the genetic syndrome of  
 RT ventricular preexcitation and conduction system disease with  
 RT childhood onset and absence of cardiac hypertrophy.";  
 RL Circulation 104:3030-3033(2001).  
 RN [8]  
 RP VARIANT'S HCM/WPWS LEU-350 INS AND ARG-383.  
 RX MEDLINE=21264334; PubMed=11371514;

RA Blair E., Redwood C., Ashrafian H., Oliveira M., Broxholme J.,  
 RA Kerr B., Salmon A., Oestman-Smith I., Watkins H.;  
 RT "Mutations in the gamma(2) subunit of AMP-activated protein kinase  
 RT cause familial hypertrophic cardiomyopathy: evidence for the central  
 RT role of energy compromise in disease pathogenesis.";  
 RL Hum. Mol. Genet. 10:1215-1220(2001).  
 RN [9]  
 RP VARIANT WPWS GLN-302.  
 RX MEDLINE=21279949; PubMed=11407343;  
 RA Gollob M.H., Green M.S., Tang A.S.-L., Gollob T., Karibe A.,  
 RA Al Sayegh A.H., Ahmad F., Lozado R., Shah G., Fananapazir L.,  
 RA Bachinski L.L., Roberts R.;  
 RT "Identification of a gene responsible for familial  
 RT Wolff-Parkinson-White syndrome.";  
 RL New Engl. J. Med. 344:1823-1831(2001).  
 RN [10]  
 RP ERRATUM.  
 RA Gollob M.H., Green M.S., Tang A.S.-L., Gollob T., Karibe A.,  
 RA Al Sayegh A.H., Ahmad F., Lozado R., Shah G., Fananapazir L.,  
 RA Bachinski L.L., Roberts R.;  
 RL New Engl. J. Med. 345:552-552(2001).  
 RN [11]  
 RP ERRATUM.  
 RA Gollob M.H., Green M.S., Tang A.S.-L., Gollob T., Karibe A.,  
 RA Al Sayegh A.H., Ahmad F., Lozado R., Shah G., Fananapazir L.,  
 RA Bachinski L.L., Roberts R.;  
 RL New Engl. J. Med. 346:300-300(2002).  
 RN [12]  
 RP VARIANT'S WPWS GLN-302; ASN-400 AND ILE-488.  
 RX MEDLINE=21686220; PubMed=11827995;  
 RA Arad M., Benson D.W., Perez-Atayde A.R., McKenna W.J., Sparks E.A.,  
 RA Kanter R.J., McGarry K., Seidman J.G., Seidman C.E.;  
 RT "Constitutively active AMP kinase mutations cause glycogen storage  
 RT disease mimicking hypertrophic cardiomyopathy.";  
 RL J. Clin. Invest. 109:357-362(2002).  
 CC -1- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.  
 CC -1- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: ISOFORM B IS UBIQUITOUSLY EXPRESSED EXCEPT IN  
 CC LIVER AND THYMUS. THE HIGHEST LEVEL IS DETECTED IN HEART WITH  
 CC ABUNDANT EXPRESSION IN PLACENTA AND TESTIS.  
 CC -1- DISEASE: Defects in PRKAG2 are the cause of Wolff-Parkinson-White  
 CC syndrome (WPWS or preexcitation syndrome). It is the second most  
 CC common cause of paroxysmal supraventricular tachycardia. WPWS can  
 CC be associated with a form of hypertrophic cardiomyopathy (HCM),  
 CC which is probably due to polysaccharide storage in the heart.  
 CC Defects in PRKAG2 may not be a frequent cause of HCM where no  
 CC features of pre-excitation are found in affected individuals.  
 CC -1- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ249976; CAB65116.1; -  
 CC EMBL: AF087875; AAK00413.1; -  
 CC EMBL: AK001887; BAA91962.1; -  
 CC EMBL: BC020540; AAH20540.1; -  
 CC EMBL: AB025580; BAA84695.1; -  
 CC EMBL: AC006966; AAF03528.1; -  
 CC Genew: iGNC:9386; PRKAG2.





[illegible]



CC	EMBL; AJ277480; CAB89520.1; -
DR	InterPro; IPR000644; CBS_domain.
DR	Pham; PF00571; CBS; 4.
DR	SMART; SM00116; CBS; 4.
KW	Carbohydrate metabolism; Transcription regulation; Nuclear protein;
FT	Repeat; CBS domain..
FT	DOMAIN 40 94 CBS 1.
FT	DOMAIN 122 180 CBS 2.
FT	DOMAIN 198 252 CBS 3.
FT	DOMAIN 271 324 CBS 4.
SQ	SEQUENCE 328 AA; 37163 MW; DC9ED3F85E46BAD3 CRC64;
	Query Match 16.5%; Score 418.5; DB 1; Length 328;
	Best Local Similarity 31.0%; Pred. No. 1.7e-24;
	Matches 99; Conservative 72; Mismatches 121; Indels 27; Gaps 6;
QY	177 RKPGAQIYRMQEHCTCDAMATSSKLIVFDTMLETKKAFFALVANGVRAAPLWDSSKKGS 236
DB	20 QKLAQSVIRVQKSTSDVPLVSYRIVLDTLSLLVKKSLNLIQNNVSAPIWDRAQKTSK 79
QY	237 FVGMLTITFDILVHRYRSPSLVQIYIEQHKIETWREIYLQGC--FKPLVSI SPNDS-- 292
DB	80 FAGLLTSSDFINVIQYFHPN-----DKFELVDKLQNLGLKDIERAIGIQPYDTRS 130
QY	293 -----LFEAVYTLKNIIRHLPLVDPVSGN-----VLHILTHKLLKFLHFGSLLPPRS 342
DB	131 IHPRPFLYACVKMIESRRPILIDQESTQREIVSVSLTYQYRILKFAVNLCKEIR--R 187
QY	343 FLYRTQIDIGIGTFDLAVVLETAPILTALDIF-VDRRVSAIPVYNECGOVGLYSRFDV 401
DB	188 YLKAPRELDIISTNNINSCMSTPVIQTLTLTLAGGVSSVPVDEQGLVNVYEAVDV 247
QY	402 IHLAAQQTNYNHLDMVSVEALRQRTLCLEGVLSQCPHESIGEVIDRIAREQVHRLVLVDET 461
DB	248 LGLIKGIYNDLSLVSGEALMRSSDDFEGVYFCTEKHLLSILTQVRSVRHFEFVVDNSN 307
QY	462 QHLLGVVVSLSDIQLQALVLS 480
DB	308 GFLTGVLTLSDILKYLFLA 326
RESULT 11	
AAKG_PIG	STANDARD; PRT; 133 AA.
ID	A09138;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE	(AMPK) (38 kDa subunit) (Fragments).
GN	PRKAG1.
OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxId-9823;
RP	{1}
RN	SEQUENCE.
RC	TISSUE=Liver;
RC	MEDLINE=95050763; PubMed=7961907;
RA	Stapleton D., Gao G., Michell B.J., Widmer J., Mitchelhill K.,
RA	Teh T., House C.M., Witters L.A., Kemp B.E.;
RT	"Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
RT	homologs of proteins that interact with yeast Snf1 protein kinase.";
RL	J. Biol. Chem. 269:29343-29346(1994).
CC	-!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC	SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC	REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC	INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC	SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC	-!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC	GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC	-!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA









Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1085.5	42.8	566	11	Q91WC5	Q91wg5 mus musculus
2	871.5	34.3	647	5	Q96613	Q96613 drosophila
3	869.5	34.3	634	5	Q9VDD2	Q9vdd2 drosophila
4	869.5	34.3	906	5	Q8SX78	Q8sxt8 drosophila
5	869.5	34.3	1400	5	Q8S257	Q8szs7 drosophila
6	736.5	22.0	372	5	Q9BHJ6	Q9bhl6 caenorhabdi
7	559.5	29.0	577	5	Q8SSV7	Q8ssv7 dictyosteli
8	488	19.2	423	5	Q02168	Q02168 caenorhabdi
9	470.5	18.5	448	5	Q9N501	Q9n501 caenorhabdi
10	374.5	14.8	487	10	Q94A46	Q94A46 arabidopsis
11	364	14.3	382	10	Q9FV59	Q9fv59 arabidopsis
12	346.5	13.7	497	10	Q9FYJ5	Q9fyj5 zea mays (m
13	345	13.6	496	10	Q9FYJ4	Q9fyf4 zea mays (m
14	243.5	9.6	391	10	Q04028	Q04028 arabidopsis
15	227	8.9	423	5	Q95ZL0	Q9szl0 caenorhabdi
16	225	8.9	63	11	Q92SV0	Q92sv0 mus musculus



```

Db 149 DLEEDDSQIFVKKFRFHKCYDLIPTSAKLIVVFDTQLLVKKAFYALVYNGVRAAPLWDSK 208
QY 235 QSFVGLMTITDFILVHRYRSPVLQIYEIQHKIETWREIYLOQCFKPLVISPNDSLF 294
Db 209 QQFVGLMTITDFIKILOMYKSPNASMEQLEEHKLDTWRSV-LHNQVMPVLSIGPDASLY 267
QY 295 EAVYTLKRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLRPSFYRTIQDLGIG 354
Db 268 DAIKILHRSRIHRLPVIDPATGNVLYILTHKRIHLRFLYINELPKPAYMQKSRLREKIG 327
QY 355 TFRDLAVVLETAPILTALDIFVDRVSALPVNNECGOVVGLYSRFDVTHLAAQOITYNHLD 414
Db 328 TYNNIETADETTSIITALKKFFVERVSALPLVDSGRLDVIYAKFDVINLAAEKTYNDLD 387
QY 415 MSVGEALRORTLCLEGVLSQPHESLGEVIDRIARQVHRLVLDVETQHLLGLVGVVLSDDL 474
Db 388 VSLRKANEHRNEFWESGVQKCNLDESPLYTIMERIVRAEVHRLVVDENRKVIGIISLSDIL 447
QY 475 QALVLSPPAG 483
Db 448 LYLVRPSPG 456

RESULT 4
Q8SXT8
ID Q8SXT8 PRELIMINARY; PRT; 906 AA.
AC Q8SXT8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RE22690p.
GN SNF4AGAMMA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY084138; AAL89876.1; -.
SQ SEQUENCE 906 AA; 99970 MW; C867D9556F42D57F CRC64;

Query Match 34.3%; Score 869.5; DB 5; Length 906;
Best Local Similarity 53.1%; Pred. No. 5e-62;
Matches 164; Conservative 73; Mismatches 71; Indels 1; Gaps 1;

QY 175 ELRKGAGIYMRFMQEHCTCYDAMATSSKIIVFTMTLEIKKAFALVANGVRAAPLWDSK 234
Db 421 DLEEDDSQIFVKKFRFHKCYDLIPTSAKLIVVFDTQLLVKKAFYALVYNGVRAAPLWDSK 480
QY 235 QSFVGLMTITDFILVHRYRSPVLQIYEIQHKIETWREIYLOQCFKPLVISPNDSLF 294
Db 481 QQFVGLMTITDFIKILOMYKSPNASMEQLEEHKLDTWRSV-LHNQVMPVLSIGPDASLY 539
QY 295 EAVYTLKRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLRPSFYRTIQDLGIG 354
Db 540 DAIKILHRSRIHRLPVIDPATGNVLYILTHKRIHLRFLYINELPKPAYMQKSRLREKIG 599
QY 355 TFRDLAVVLETAPILTALDIFVDRVSALPVNNECGOVVGLYSRFDVTHLAAQOITYNHLD 414
Db 600 TYNNIETADETTSIITALKKFFVERVSALPLVDSGRLDVIYAKFDVINLAAEKTYNDLD 659
QY 415 MSVGEALRORTLCLEGVLSQPHESLGEVIDRIARQVHRLVLDVETQHLLGLVGVVLSDDL 474
Db 660 VSLRKANEHRNEFWESGVQKCNLDESPLYTIMERIVRAEVHRLVVDENRKVIGIISLSDIL 719

```

```

QY 475 QALVLSPPAG 483
Db 720 LYLVRPSPG 728

RESULT 5
Q8SZS7
ID Q8SZS7 PRELIMINARY; PRT; 1400 AA.
AC Q8SZS7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LD22662p.
GN SNF4AGAMMA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY070541; AAL48012.1; -.
SQ SEQUENCE 1400 AA; 152380 MW; 411B93CC6B9EC7AF CRC64;

Query Match 34.3%; Score 869.5; DB 5; Length 1400;
Best Local Similarity 53.1%; Pred. No. 9e-62;
Matches 164; Conservative 73; Mismatches 71; Indels 1; Gaps 1;

QY 175 ELRKGAGIYMRFMQEHCTCYDAMATSSKIIVFTMTLEIKKAFALVANGVRAAPLWDSK 234
Db 915 DLEEDDSQIFVKKFRFHKCYDLIPTSAKLIVVFDTQLLVKKAFYALVYNGVRAAPLWDSK 974
QY 235 QSFVGLMTITDFILVHRYRSPVLQIYEIQHKIETWREIYLOQCFKPLVISPNDSLF 294
Db 975 QQFVGLMTITDFIKILOMYKSPNASMEQLEEHKLDTWRSV-LHNQVMPVLSIGPDASLY 1033
QY 295 EAVYTLKRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLRPSFYRTIQDLGIG 354
Db 1034 DAIKILHRSRIHRLPVIDPATGNVLYILTHKRIHLRFLYINELPKPAYMQKSRLREKIG 1093
QY 355 TFRDLAVVLETAPILTALDIFVDRVSALPVNNECGOVVGLYSRFDVTHLAAQOITYNHLD 414
Db 1094 TYNNIETADETTSIITALKKFFVERVSALPLVDSGRLDVIYAKFDVINLAAEKTYNDLD 1153
QY 415 MSVGEALRORTLCLEGVLSQPHESLGEVIDRIARQVHRLVLDVETQHLLGLVGVVLSDDL 474
Db 1154 VSLRKANEHRNEFWESGVQKCNLDESPLYTIMERIVRAEVHRLVVDENRKVIGIISLSDIL 1213
QY 475 QALVLSPPAG 483
Db 1214 LYLVRPSPG 1222

RESULT 6
Q9BHL6
ID Q9BHL6 PRELIMINARY; PRT; 372 AA.
AC Q9BHL6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Y11B2A.8 protein.
GN Y11B2A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```



```

Query Match          19.2%; Score 488; DB 5; Length 423;
Best Local Similarity 36.7%; Pred. No. 2e-31;
Matches 119; Conservative 69; Mismatches 106; Indels 30; Gaps 8;

QY 182 QIWRPMQEHCTYDAMATSSKLVIFDTMLKKAFFALVANGVRAAPLWDSK-KQSFVGM 240
Db 38 EAFARLLWLNQCYEAMPSSKVVFDQGLLHKAFNGLLAQSTRHVLSSDPDFGKLDGI 97

QY 241 LTTDTFLVLHRYYS-----PLVQI--YIEQHKIETWRE--IYLOGCFKPLVSI 286
Db 98 LSVTDFIKVMLKIYERTCKCKESTELDMTQIANEEIGNLSIRQYRELVKKEGNRLPLVS 157

QY 287 ISPNDLSFEAVYTLKRNHRLPVLDPVSGNVLIHTHKLKLFHIFGSLLPRESPLYR 346
Db 158 VDASSGLLDAACILAEHRVHRIPVDPDGSALFILTHRIKLFKWLFGKHLAPLEYLKH 217

QY 347 TIODLGIGTRDLAVVLETAPILTALDIFVDRVRSALPVV-NECQVVVGLYSRFDVTHLA 405
Db 218 SPKELGIGTWSGIRVFPDTQLVDCLDILLKGVSGLPVVERETFKVVDMYSRFDVAVGIA 277

QY 406 AQOTYHNLDMVGEALRORTLCLEG-----VLSQPHESLGEVIDRTAREOVHRLVL 457
Db 278 LE---NKLDTIKVKEALAFKS---QGGPMKNDERVWSVRDNEFMRKAVNLVDHNVHRLCA 331

QY 458 VDETQHLGLVWSLSDILOALVLSL 481
Db 332 VNEHGGIEGVLSVDVNFVNVQP 355

RESULT 9
Q9N501 PRELIMINARY; PRT; 448 AA.
AC Q9N501:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 51.5 kDa protein.
GN Y41G9A.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Ryan E., Wohlman P., Walker C., Fielder T.;
RT "The sequence of C. elegans cosmid Y41G9A.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006761; AAP60550.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 3.
KW Hypothetical protein.
SQ SEQUENCE 448 AA; 51514 MW; B689218979299479 CRC64;

Query Match          18.5%; Score 470.5; DB 5; Length 448;
Best Local Similarity 33.0%; Pred. No. 5.9e-30;
Matches 106; Conservative 79; Mismatches 111; Indels 25; Gaps 7;

183 IYMRPMQEHCTYDAMATSSKLVIFDTMLKKAFFALVANGVRAAPLWDSKQSFVGM 242
Db 117 VYTHLLOLSQCYEAMPARNKLVFTNDISVRKAENGLIYNCRMTGLVADSQTLEITGVLS 176

QY 243 ITDTFLVLHRY--YRSLPVOI-----YIEBQHKIETWRE--IYLOGCFKPLVSI 287
Db 177 VTDFITVIMLMLWKYRENLDLKGTPLSHEDFQMDIAYMFISSWKGLCTKGQLKPFINI 236

QY 288 SPNDLSFEAVYTLKRNHRLPVLDPVSGNVLIHTHKLKLFHIFGSLLPRESPLYRT 347
Db 237 GLKESITFRAVELITKYRIHRLPVMDEKTCDCAYILTHRRILHYIWKHCALLPKPECLSQ 296

QY 348 IODLGIGTRDLAVVLETAPILTALDIFVDRVRSALPVNE-CGVVGLYSRFDVHLLA 406
Db 297 VVDLEISGNLWLFANEQTPLECLMDLNNISGIPVOKTKLVLEVYIRFDA---AS 353

QY 407 QOTYHNLDMVSG--EALRQRT----LCLEGVLSQPHESLGEVIDRIAREOVHRLVLDE 460
Db 354 AAFSDHIDLVSVTRAIQRYQNGIRRDGVVYNTYTTLLWSLIEIFIDKKNVHRIEWD 413

QY 461 TQHLGLVWSLSDILOALVLSL 481
Db 414 RTILKGIISLSDVIEFLVLRP 434

RESULT 10
Q944A6 PRELIMINARY; PRT; 487 AA.
AC Q944A6:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Atlg09020/F7G19.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pam P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439826; AA127498.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
SQ SEQUENCE 487 AA; 53466 MW; 4545FE3BF2C4EBEC CRC64;

Query Match          14.8%; Score 374.5; DB 10; Length 487;
Best Local Similarity 29.0%; Pred. No. 4.3e-22;
Matches 120; Conservative 77; Mismatches 140; Indels 77; Gaps 17;

QY 112 TCWDCPLPSDCTASAGSS---TDDVELATEPFAPEAWECELEGLLEERPALCLSPQAPFP 168
Db 102 TGPDMVPAGFSPTLGRSNMVDVDFLTADPDSQA-----VP 139

QY 169 KL-GWDDLRKPGAQIYMRPMQEHCTYDAMATSSKLVIFDTMLKKAFFALVANGVRAA 227
Db 140 RMSGVDLELSRHRISV---LLSTRYATPELLPESGKVIADVNLVVKQAFHILYEQGIPLA 196

QY 228 PLWDSKQSFVGM 240
Db 197 PLWDFGKGQFVGLPDLFILRELPHGNSLITE-DELETHITAAWKEGKAHSRQYDG 255

QY 282 -----RPLVSI SPNDLSFEAVYTLKRNHRLPVL-----DFVSNVLIHTHKLK 329

```

```
Db 256 SGRPYRPLVQGPYDNLKQVAKILQNKVAAVPVIYSSLOGSGYPQLLHLASLSGLK 315
QY 330 ----FLHIFGSLPRPFLYRTIQDLGIGTF-----RDLAVVLETAPILTALDIFV 376
Db 316 ICRYFPHSSSL---PILOQPCISPLCTWVPRIIGESSKPLATLRPHASLGSLALV 371
QY 377 DRVYSALPVVNECGVGLYSRFDVHLAAQOTYN--HL-DMSVGEALR--QRTLCLEGV 431
Db 372 QAEVSSIPVDDNSLIDIVSRSDITALAKDKAYAQIHLDDMTVHQALQLGODASPPYGI 431
QY 432 LS-----COPHESLGEVIDRIAREQVHRLVLVDE--TOHLLGVVSLSDILQALV 478
Db 432 FNGQRCHMCLRSDSLVKVMEKLANPGVRRLLVIVEAGSKRVEGIISLSDVQFELL 485

RESULT 11
Q9FV59 PRELIMINARY; PRT; 382 AA.
AC Q9FV59;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE putative activator subunit of SNF1-related protein kinase SNF4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20387008; PubMed=10929106;
RA Kleinow T., Bhale Rao R., Breuer F., Umeda M., Salchert K., Koncz C.;
RT "Functional identification of an arabidopsis snf4 ortholog by
RT screening for heterologous multicopy suppressors of snf4 deficiency in
RT yeast.";
RL Plant J. 23:115-122(2000).
DR EMBL: AF250335; AAG10141.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 3.
KW Kinase.
SQ SEQUENCE 382 AA; 41800 MW; A9E7A4D5E1A3CB53 CRC64;

Query Match 14.3%; Score 364; DB 10; Length 382;
Best Local Similarity 28.6%; Pred. No. 2.2e-21;
Matches 118; Conservative 76; Mismatches 140; Indels 78; Gaps 17;

QY 111 PTWCDCLPDCTASAGSSHHVVELATFPATPAWECHLRLGLEERFALCLSPQAPFKI 170
Db 3 PAGF-----SPEFLGRNMDVGVPLKTADPSQFA-----VPRM 36

QY 171 -GWDELRRKQAIYRMFOHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPL 229
Db 37 SGVDLSELSHRISV---LLSTRYAYELLPESGKVIALDVNLPVKAQFAHILYEGIPLAPL 93

QY 230 WDSKKGSFVGMJLTIDFELVILHRY--YRSPVQVIEIQHKIETWRE--IYLGQCF---- 281
Db 94 WDRGKQGVGLVPLDFLLRLGLTGNSNLT-ELETHTTAAWKEGAHISKQYDQSG 152

QY 282 ----KPLVSIENSPNLSPEAVYTLIKNRJHRLPVL-----DPVSGNVHLTHKRLK--- 329
Db 153 RPYRPLVQGVGYDNLKQVAKILQNKVAAVPVIYSSIQDGSYPQLLHLASLSGLKCTC 212

QY 330 --FLHIFGSLPRPFLYRTIQDLGIGTF-----RDLAVVLETAPILTALDIFVDR 378
Db 213 RYFRHSSSL-----PILOQPCISPLCTWVPRIIGESSKPLATLRPHASLGSLALVQA 268

QY 379 RVSAIPVNECGVGLYSRFDVHLAAQOTYN--HL-DMSVGEALR--QRTLCLEGVLS 433
Db 269 EVSSIPVDDNSLIDIVSRSDITALAKDKAYAQIHLDDMTVHQALQLGODASPPYGI 428

QY 434 -----COPHESLGEVIDRIAREQVHRLVLVDE--TOHLLGVVSLSDILQALV 478
```

```
Db 329 GQRCHMCLRSDSLVKVMEKLANPGVRRLLVIVEAGSKRVEGIISLSDVQFELL 380

RESULT 12
Q9FUY5 PRELIMINARY; PRT; 497 AA.
AC Q9FUY5;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE protein kinase AKINbeta gamma-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Victoria L., Mar A., Tatjana K., Csaba K., Montserrat P.;
RT "Domain fusion between Snf1 related kinase subunits during plant
RT evolution.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF276085; AAG31751.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 3.
DR SMART: SM00116; CBS; 3.
KW Kinase.
SQ SEQUENCE 497 AA; 54938 MW; B9EF06A530C1AD8B CRC64;

Query Match 13.7%; Score 346.5; DB 10; Length 497;
Best Local Similarity 29.0%; Pred. No. 8.5e-20;
Matches 99; Conservative 72; Mismatches 109; Indels 61; Gaps 12;

QY 187 FMOHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPLWDSKKQSPVGMJRTIDF 246
Db 167 YLNLFHCYDLLPDSCKVIALDINLPVKQSFHILHEGIPVAPLWDSRFGVGLLSPLDF 226

QY 247 ILVLHRY--YRSPVQVIEIQHKIETWREIYLQCG-----FKPLVSIENSPNLSFE 295
Db 227 ILILRELETHGNSNLT-EQLETHTTISAKKEAKRQTCGRNDGQWRAHQHVLHATPYESLRD 285

QY 296 AVYTLIKNRJHRLPVLDPVSGN-----VLHILTHKRLK-----FLHIFGSL----- 337
Db 286 IAVKLLINDISTVPVIYSSSDSGSFQPLLHLASLSGLKICFRYFKNSTGNLPILNQPYC 345

QY 338 -LRPSPFLYRTIQDLGIGTFERDLAVVLETAPILTALDIFVDRVYSALPVVNECGVGLY 396
Db 346 SIPLGNSW---VPKIGDPNSRPLAMLRPNASLSALNMLVQAGVSSIPVDENDLSLDY 401

QY 397 SRFDVHLAAQOTYNHL---DMSVGEALR-----QR-TLCLEGVLSQCPHE 438
Db 402 SRSDITALAKAKVYTHVRIDENAIHQALQLGQDANTPFGFFNGRCQCLRS-----D 454

QY 439 SLGEVIDRIAREQVHRLVLVDE--TOHLLGVVSLSDILQALV 478
Db 455 PLLKVMEKLANPGVRRVIVEAGSKRVEGIISLSDIFKELL 495

RESULT 13
Q9FUY4 PRELIMINARY; PRT; 496 AA.
AC Q9FUY4;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Protein kinase AKINbeta gamma-2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
```





```
QY 216 FFALVANGVRAAPLWDSKKQSFVGKLT-----ITDFILVL--HRYRSPPLVOIVEIEOHKI 269
Db 90 VHALSQGHIAAVVNTDKYQAEVFMNGHCLTAILLVAAGNREVASKTL-----V 140
QY 270 ETWREIYLQGCCKPIVSI SPNDSLFEAVYTLIKNRIHRLPVL D---PVSGNVLHILTHKR 326
Db 141 EFLKEIGSGN-----IICSGVNSWEAANIISHNKISFVPIDTIIIPKPGTPLYFTPRM 196
QY 327 LLK-----FIIHFGSLLPRPSFLYRTIQDLGIGTFRD--LAVVLETAPILTA 371
Db 197 ILQETVLKLSDFGDAILLHV-----RQATLDQKKIGTWNDDVLKIGLNTT-IEEA 245
QY 372 LDIFVDRRVSALPVVNECCQVVGLYSRFDVI-HLAAQQTYNHLDMSVGEALRORTLCLEG 430
Db 246 IKLMSERKMSITIPVVNDEFKQIVNMLARKDIIILEIMSHQGGNFHDM-----LKEPVKILQS 300
QY 431 VLS---COPHESLGEVIDRIAREQVHRLVLDVETQHLIGVVVSLSDIL 474
Db 301 LQSRLYGRSSYTVFETVAKMWTSDKSSLPIDEGKRILAVVSCSDIL 348
```

Search completed: June 6, 2003, 11:02:49  
Job time : 66.373 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 29.1173 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2538  
Sequence: 1 MBPGLHEALRTPSWSSLG.....LSDIQLVALSPAGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1080.5	42.6	328	2	US-08-878-989-7
2	1080.5	42.6	328	4	US-09-272-796-7
3	1028	40.5	330	4	US-09-359-161-6
4	1017	40.1	331	2	US-08-878-989-21
5	1017	40.1	331	3	US-09-101-146-64
6	1017	40.1	331	4	US-09-272-796-21
7	457	18.0	322	4	US-09-359-161-7
8	214	8.4	379	4	US-09-359-161-5
9	190	7.5	373	4	US-09-359-161-3
10	102	4.0	2273	4	US-09-426-998-5
11	101	4.0	187	4	US-09-199-637A-287
12	97	3.8	852	2	US-09-070-060-3
13	97	3.8	852	3	US-09-357-746-3
14	96	3.8	830	4	US-09-562-737-33
15	94.5	3.7	1285	3	US-07-582-945-2
16	94.5	3.7	1285	2	US-08-453-141-2
17	94.5	3.7	1285	3	US-08-293-314-2
18	94	3.7	830	4	US-09-562-737-37
19	94	3.7	2353	4	US-08-984-709A-50
20	93.5	3.7	1208	4	US-09-463-702A-2
21	91.5	3.6	854	3	US-09-070-060-4
22	91.5	3.6	854	3	US-09-357-746-4
23	91.5	3.6	2972	4	US-09-579-181-2
24	91.5	3.6	3118	4	US-09-379-181-1
25	90.5	3.6	443	4	US-09-134-001C-3767
26	90.5	3.6	599	2	US-08-426-125-1
27	90.5	3.6	599	2	US-08-455-355-1

28 90.5 3.6 909 4 US-09-013-895A-4 Sequence 4, Appli  
29 90.5 3.6 909 4 US-09-448-868-4 Sequence 4, Appli  
30 1618 1 3.6 1618 1 US-07-853-913-4 Sequence 4, Appli  
31 90.5 3.6 2016 4 US-09-634-920-4 Sequence 4, Appli  
32 90 3.5 595 4 US-08-764-870-12 Sequence 12, Appli  
33 90 3.5 595 4 US-08-980-115-12 Sequence 12, Appli  
34 90 3.5 619 1 US-08-465-746-2 Sequence 2, Appli  
35 90 3.5 619 1 US-08-214-164-2 Sequence 2, Appli  
36 90 3.5 619 2 US-08-467-852A-3 Sequence 3, Appli  
37 90 3.5 619 2 US-08-246-636-2 Sequence 3, Appli  
38 90 3.5 619 2 US-08-247-491A-3 Sequence 3, Appli  
39 90 3.5 619 2 US-08-319-795-2 Sequence 2, Appli  
40 90 3.5 619 2 US-08-468-985-2 Sequence 2, Appli  
41 90 3.5 619 3 US-08-312-949-2 Sequence 2, Appli  
42 90 3.5 648 1 US-08-072-070-2 Sequence 2, Appli  
43 90 3.5 648 1 US-08-469-434-2 Sequence 2, Appli  
44 90 3.5 648 1 US-08-214-222-2 Sequence 2, Appli  
45 90 3.5 648 2 US-08-467-852A-2 Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-878-989-7  
; Sequence 7, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PENITUT01



QY 165 APPPKLWGDDELKPGAGQIYRMFQEHCTCYDAMATSSKLVFDTLMLEIKKAFKAFALVANGV 224  
Db 9 APAPNEHSQETPESNSVVTETKSHRCYDLIPTSSKLVVFDTSIQVKKAFKAFALVTNGV 68  
QY 225 RAAPLWDSKKSFGVGMITITDFILVHRYRSPVLOIYEIEQHKIETWREIYVLCQCFKPL 284  
Db 69 RAAPLWDSKKSFGVGMITITDFINILHRYKXSALVQIYELEEHKIEIETWREIYVLCQCFKPL 128  
QY 285 VTSIPNDSLEFVAVTILKNIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPPSPFL 344  
Db 129 VCLSPNASLFAVSSLRKNIHRLPVPDPSGNTLIHLTHKRLKFLHIFGSLPPSPFL 188  
QY 345 YRTIODLGIGTFRLAVLVTAPILFALDIFVDRRVSAIPVNVCGQVGLYSRFDVHL 404  
Db 189 SKSLEELQIGTYANIAMVTTTPVYVVALGIFVQHRVSAIPVNVCGQVGLYSRFDVHL 248  
QY 405 AAQOTYHNLDSVGEALRQRTLCLEGLVSCQPHESLGEVIDRTAREQVHRLVLDVETQHL 464  
Db 249 AAQOTYHNLDSVGEALRQRTLCLEGLVSCQPHESLGEVIDRTAREQVHRLVLDVETQHL 308  
QY 465 LGVYSLDILQALVLS 480  
Db 309 KGIVSLDILQALVLT 324

RESULT 4

US-08-878-989-21  
; Sequence 21, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 331 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1335856  
US-08-878-989-21  
Query Match 40.1%; Score 1017; DB 2; Length 331;  
Best Local Similarity 64.8%; Pred. No. 2.7e-104;  
Matches 193; Conservative 53; Mismatches 52; Indels 0; Gaps 0;  
QY 183 IYMRPQEHCTCYDAMATSSKLVFDTLMLEIKKAFKAFALVANGVRAAPLWDSKKSFGVGMILT 242  
Db 28 VYTFPKSHRCYDLIPTSSKLVVFDTSIQVKKAFKAFALVTNGVRAAPLWDSKKSFGVGMILT 87  
QY 243 ITDFILVHRYRSPVLOIYEIEQHKIETWREIYVLCQCFKPLVSPNDSLEFVAVTILK 302  
Db 88 ITDFINILHRYKXSALVQIYELEEHKIEIETWREIYVLCQCFKPLVSPNDSLEFVAVTILK 147  
QY 303 NRHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPPSPFLYRTIODLGIGTFRLAVV 362  
Db 148 NKHRLPVIDPESGNTLIHLTHKRLKFLHIFGSLPPSPFLYRTIODLGIGTFRLAVV 207  
QY 363 LETAPILTALDIFVDRRVSAIPVNVCGQVGLYSRFDVHLAAQOTYHNLDSVGEALR 422  
Db 208 RTTPVYVVALGIFVQHRVSAIPVNVCGQVGLYSRFDVHLAAQOTYHNLDSVGEALR 267  
QY 423 QRTLCLEGLVSCQPHESLGEVIDRTAREQVHRLVLDVETQHLGLWVSLDILQALVLS 480  
Db 268 HRSHYEGVGLKCYLHETLEIINRLVEAEVHRLVVDENDVVKGVSLDILQALVLT 325  
RESULT 5  
US-09-101-146-64  
; Sequence 64, Application US/09101146  
; Patent No. 6124125  
; GENERAL INFORMATION:  
; APPLICANT: Dartmouth College, St. Vincents Institute of  
; APPLICANT: Medical Research, Kemp et al.  
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: 66 E. Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: WINDOWS 95  
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/101,146  
; FILING DATE: October 7, 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PN7450  
; FILING DATE: 8 JAN 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: DG-0050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (856) 810-1515  
; TELEFAX: (856) 810-1454  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 331  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-09-101-146-64  
Query Match 40.1%; Score 1017; DB 3; Length 331;



QY 350 DLGIGTFRDLAVVLETAPILTALDIFVDRVRVSALPVPVNECGQVGLYSRFDVHIAAAQQT 409  
Db 190 DNLIIQDNMKSCQMTTVIDVIOQLMTQGRVSSVPIIDENGYLINVEAYDVGLGKGGI 249  
QY 410 YNHLDSVGEALRQRTCLGVLSCOPHESLGEVIDRIAREOVHRLVLVDQTHLLGVYS 469  
Db 250 YNLSLSVGEALMRRSDDFEGVYCTTNDKLSITMDNIKRARVHRFFVVDVGVGLVGLT 309  
QY 470 LSDILOALVL 479  
Db 310 LSDILKYILL 319  
RESULT 8  
US-09-359-161-5  
; Sequence 5, Application US/09359161A  
; Patent No. 6342656  
; GENERAL INFORMATION:  
; APPLICANT: Bradford, Kent J.  
; APPLICANT: Dahal, Peetambar  
; APPLICANT: Yang, Hong  
; APPLICANT: Cooley, Michael  
; APPLICANT: Downie, Bruce  
; APPLICANT: Gee, Oliver  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses  
; FILE REFERENCE: 023070-095900US  
; CURRENT APPLICATION NUMBER: US/09/359,161A  
; CURRENT FILING DATE: 1999-07-21  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Phaseolus vulgaris  
; FEATURE:  
; OTHER INFORMATION: Phaseolus vulgaris L. Pv42  
US-09-359-161-5  
Query Match 8.4%; Score 214; DB 4; Length 379;  
Best Local Similarity 23.5%; Pred. No. 6.4e-15;  
Matches 85; Conservative 73; Mismatches 122; Indels 82; Gaps 17;  
QY 188 MQEHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPL-----W-----DS 232  
Db 17 LKEKKVDMWCKKRLVEVPYTSALQIMNTLVANKIVAVPVAAPPGOWICAGGSMIVES 76  
QY 233 KKQS-----FVGMLTITDFILVL-----HRYKSPVLQYIETEIQHKIETWREIY---LOG 279  
Db 77 DKQTGAVRKHYIGVMTMLDILAHITAGDDHLSCGDNITQ--DLDQMSDSVSSIIGHSPFG 134  
QY 280 CFKPLVSIISPNDLSFEAVYTLIKNRIHRLPVLDPVSG---NV-----LHILH 324  
Db 135 L--SLMTLAPNTSLMDCNEVFSKG--VHRAMV---PVDGLEENVAGSVELTESASSYQMLTQ 189  
QY 325 KRLKLFHIFGSLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRVRVSALP 384  
Db 190 MDMLKFLHGGGAEL--HSILSRVQDLGADTVQIYAIT-DRTKLVHAIKLCRAMLNAV 246  
QY 385 VW-----NECGQVGLYSRFDV--IHIAAQQTYNHLD-MSVGEALRQRTL 426  
Db 247 IVRATGVGDHDKQLNGRCRKLIGTFSATDLRGCHISSLSKSLWLGISALAFTEVRSSPL 306  
QY 427 CLEG-----VLSQPHESLGEVIDRIAREOVHRLVLVDQTHLLGVYSLSILQ 475  
Db 307 YSEDMONRGSRRELVTCTYAESPLSEVIEKAVTSHVRVWVYDQGLGVVSLTDVIR 366  
QY 476 AL 477  
Db 367 VI 368

RESULT 9  
US-09-359-161-3  
; Sequence 3, Application US/09359161A  
; Patent No. 6342656  
; GENERAL INFORMATION:  
; APPLICANT: Bradford, Kent J.  
; APPLICANT: Dahal, Peetambar  
; APPLICANT: Yang, Hong  
; APPLICANT: Cooley, Michael  
; APPLICANT: Downie, Bruce  
; APPLICANT: Gee, Oliver  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses  
; FILE REFERENCE: 023070-095900US  
; CURRENT APPLICATION NUMBER: US/09/359,161A  
; CURRENT FILING DATE: 1999-07-21  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
; FEATURE:  
; OTHER INFORMATION: Lycopersicon esculentum sucrose non-fermenting protein  
; OTHER INFORMATION: Kinase activation subunit 4 (LeSNF4)  
US-09-359-161-3  
Query Match 7.5%; Score 190; DB 4; Length 373;  
Best Local Similarity 22.4%; Pred. No. 2.9e-12;  
Matches 82; Conservative 66; Mismatches 126; Indels 92; Gaps 15;  
QY 186 RFMOHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPL-----W-----230  
Db 20 QMLADKQVKDLIIDKRLVEVPYATADTINILMANKVVAVPVAAPPGHIGAGGSNIL 79  
QY 231 DSKQS-----FVGMLTITDFI-LVLHRYRSP-----LVQIYEIQHKIETWREI 275  
Db 80 ESDKQTGAVRKHYIGVMTMLDILAVIAGNGYRDDDDDLTKMWPVPSIIIGHCLES----135  
QY 276 YLQCFKPLVSIISPNDLSFEAVYTLIKNRIHRLPVLDPVSG---NVL-----HI 321  
Db 136 -----LSLMTLSPNTSIVDCNEVFSKG--IHRAMV---PVNGLEENVAGSVELTESASCYRM 186  
QY 322 LTHKRLKLFHIFGSLPRPSFLYRTIQDLGIGTFRDLAV-VLETAPILTALDIFVDRVR 380  
Db 187 LTQMDLLRFLNDQOEL---KAIMSHKYSDDKQLOAITDTVFGVTNKAKVIDIKCMRTASL 243  
QY 381 SALPVPVNECG-----QVVGYSRFDVHIAAAQQTYNHLDMSVGEALRQRT 425  
Db 244 NAVPIVSSNDITEDHTQLVNGKKRKIVGTFSATDLRGCPVKMQPPLLNLVLDLAKM--301  
QY 426 LCLEGV-----LSCOPHESLGEVIDRIAREOVHRLVLVDQTHLLGVYSLSDI 473  
Db 302 --LSGAPNTGLRSWREQVTCRPSSLGEVVEKYSVDSNVHRVWVYDQGLGVVSLTDM 359  
QY 474 LOALVL 479  
Db 360 IRVIRL 365  
RESULT 10  
US-09-426-998-5  
; Sequence 5, Application US/09426998  
; Patent No. 6358706  
; GENERAL INFORMATION:  
; APPLICANT: DUBIN, ADRIENNE E.  
; APPLICANT: PYATI, JAYASHREE  
; APPLICANT: ZHU, JESSICA Y  
; APPLICANT: ERLANDER, MARK G  
; APPLICANT: GALINDO, JOSE E  
; TITLE OF INVENTION: DNA ENCODING HUMAN ALPHAIG T-TYPE CALCIUM  
; CHANNEL (ALPHAIG-C)

FILE REFERENCE: ORT-1057  
CURRENT APPLICATION NUMBER: US/09/426,998  
CURRENT FILING DATE: 1999-10-26  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PATENTIN VER. 2.0  
SEQ ID NO 5  
LENGTH: 2273  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-09-426-998-5

Query Match 4.0%; Score 102; DB 4; Length 2273;  
Best Local Similarity 19.5%; Pred. No. 0.4;  
Matches 122; Conservative 77; Mismatches 194; Indels 234; Gaps 30;

QY 2 EPGLHAKRTSWSLGGSHQMGFLEQENSSSSPSPAVTSSSERIRGKRAKALRW 61  
DB 1076 EPGAHEMKSPSARS---SPHSWS-----RASSWTS-----RRSSNSLG 1114

QY 62 ROKSVEEGEPGQ-----GEGPRSRAAESTGLEATFPKTTPLAQADPAGVGTPTG-- 113  
DB 1115 RAPSLKRRSPGERSLLSGEQSQDQESSEEE-----RASPAGSDHRRHGS 1164

QY 114 -----WCLPSDC-----TASAAGSSTDVE-----LATEPPATEAWE 146  
DB 1165 EREAKSSFD-LPDTLOVPLGHTASGRGSAHEQDNGKASGRLARALRFPDPLDGD 1223

QY 147 CEGLGLE--ER-----PALCLSPQAPFPKLGWDELKPKGA--QIYMFQMOETC 193  
DB 1224 ADDEGNLSGKERVRAWIRARLPACCLERDS-----WSAYIFPPQSRELLCHRIITHKM 1277

QY 194 YDAMATSSKLVTFTMLEI-----KKAFTALVANGVRAAPL-----W 230  
DB 1278 FDHWL---VIFLNCITIAMERPKIDPHSAERIFLTLSNVIFTAVFLAEMTVKVVALGW 1334

QY 231 DSKQSFV-GMLTIDFTLVHLRYRSPVQIYEIQHKIETWETIYLCQCFPLVISIP 289  
DB 1335 CFGEQAYLRSSNNWLDGLLVLSVIDILVSWSDSCTKILGMLRVLRLLRLTLRPLRVISR 1394

QY 290 NDSLFEAVYTLKRIHRLPLDVPVSGNVHLHILTKRLKFLHIFGSL--LPRPSFL-- 344  
DB 1395 AQGLKLVETLSS-----LKPI-GNIVI-----CAFFIIFGILGQLEKGFVC 1441

QY 345 -----YRTQDLGIGTFRDLAVVLETAPILTALDIFVDRVRSALPV 386  
DB 1442 QGEDTRNITKSDCAEASRYRVRHK--YNFQNLGQALMSLEVLASKDQWVDI----- 1491

QY 387 NECCQVVGILYRFDVIHIAAQTYNH-----LMSVGEALRORTLC 427  
DB 1492 -----MYDGLDAGVDQOPINHNPMWMLLYFTSLLIYAVFVLNMFV-----VV 1536

QY 428 LEGVLSOPHESLGVE-----IDRIAREQ-----VHRL 455  
DB 1537 VENFKRCRQHEEAREERKRLRLEKKRSKQMAEQCKPYYSIDSRFLVHHL 1596

QY 456 VLVDETQHL-----GVVVSUJLQAL 477  
DB 1597 C-----TSHVLDLFTIGVLGNVVTMAM 1619

RESULT 11  
US-09-199-637A-287  
Sequence 287, Application US/09199637A  
Patent No. 6355411  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick  
APPLICANT: Goodman, Howard M.  
APPLICANT: Rahme, Laurence G.  
APPLICANT: Mahajan-Miklos, Shalina  
APPLICANT: Tan, Man-Wah  
APPLICANT: Cao, Hui  
APPLICANT: Drenkard, Eliana  
APPLICANT: Tsongalis, John

TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
FILE REFERENCE: 00786/361002  
CURRENT APPLICATION NUMBER: US/09/199,637A  
CURRENT FILING DATE: 1998-11-25  
PRIOR APPLICATION NUMBER: 60/066,517  
PRIOR FILING DATE: 1997-11-25  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 287  
LENGTH: 187  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-199-637A-287

Query Match 4.0%; Score 101; DB 4; Length 187;  
Best Local Similarity 25.0%; Pred. No. 0.0069;  
Matches 44; Conservative 14; Mismatches 46; Indels 72; Gaps 10;

QY 38 PSPAVTSS-----SERINGKRRAKALRWTRQ-----KSVEEGEPGQ 75  
DB 19 PACPTTSCPPRSRGRCSTRSAGSDR-KGRRANGARRWTRFLPPRGRSLADAAPCA 77

QY 76 EGPRSRAAEST-GLEAT-----FPKTTPLAQAD-----PAGVGTPTGWD---C 116  
DB 78 AASGRAPRASSTPGYRSTWKPLRRFRRSPCCAATGSRAGRPSAGCVPGFWSGTGC 137

QY 117 LPSD-----CTASAGSSTDDELATEPPATEAWECEGLELLEERPALCLSPQAPFP 168  
DB 138 APCSRGRPCAAP-----PPSPAW-----PARSSAGSAPSP 167

RESULT 12  
US-09-070-060-3  
Sequence 3, Application US/09070060  
Patent No. 5976849  
GENERAL INFORMATION:  
APPLICANT: Hustad, Carolyn M.  
APPLICANT: Ghildyal, Namit  
TITLE OF INVENTION: Human E3 Ubiquitin Protein  
TITLE OF INVENTION: Ligase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ZENCA Pharmaceuticals, Inc.  
STREET: 1800 Concord Pike  
CITY: Wilmington  
STATE: DE  
COUNTRY: USA  
ZIP: 19850-3437  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070,060  
FILING DATE: 30-APR-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/073,839  
FILING DATE: 05-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Higgins, Patrick H  
REGISTRATION NUMBER: 39,709  
REFERENCE/DOCKET NUMBER: PHM.70312  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302.886.4889  
TELEFAX: 302.886.8221  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 852 amino acids  
TYPE: amino acid



STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-070-060-3

Query Match 3.8%; Score 97; DB 2; Length 852;  
Best Local Similarity 22.1%; Pred. No. 0.26;  
Matches 64; Conservative 40; Mismatches 114; Indels 72; Gaps 15;  
  
QY 29 LEQNSSWSPSPAVTSSSRIRGKRAKALRWTRKSVKEGEPGQGEGR-----SRPAA 84  
DB 134 LESEVVTNGETTCSESQNDGSRKSDTRVSTNGS--DDPEDAGAGNRVSGNNSPSL 192  
  
QY 85 ESTGLEATFP-----KTTPLAQADPAGVGTPTGWCDCPLSDCTASAAGS-----STDDVE 134  
DB 193 SNGGFKPSRPPRPPTPRPASVNGSPSA-----TSEDGSGSTGSLPPTNTNINTE 248  
  
QY 135 LATE---FPATEMECELEGLLEERPALCLSP--QAPFPKLGWDELKPKCAQIYMRMQ 189  
DB 249 GATSGLIPLT-----ISGSGRPP---LNPVTOAPLPP--GWQRVDQHGRIYYVDHVE 298  
  
QY 190 EHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPL---WDSKKQSFVGMILTIDF 246  
DB 299 KRTTWD-----RPELPFGWERRVDN--MGR1YYVDH 328  
  
QY 247 ILVLRHYRSPLVQIYEIQHK1ETWREIYVLCQCFPLVS---ISPNDSLF 294  
DB 329 FTRTTTQWPTLESVRNYEQWLQ---RSQLOGAMQGFNRFIYGNODLF 375

RESULT 13  
US-09-357-746-3  
; Sequence 3, Application US/09357746  
; Patent No. 6087122  
; GENERAL INFORMATION:  
; APPLICANT: ZENCA Limited  
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE  
; FILE REFERENCE: PHM.70312.N1  
; CURRENT APPLICATION NUMBER: US/09/357,746  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: US No. 608712209/070.060  
; EARLIER FILING DATE: 1998-04-30  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 852  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-357-746-3

Query Match 3.8%; Score 97; DB 3; Length 852;  
Best Local Similarity 22.1%; Pred. No. 0.26;  
Matches 64; Conservative 40; Mismatches 114; Indels 72; Gaps 15;  
  
QY 29 LEQNSSWSPSPAVTSSSRIRGKRAKALRWTRKSVKEGEPGQGEGR-----SRPAA 84  
DB 134 LESEVVTNGETTCSESQNDGSRKSDTRVSTNGS--DDPEDAGAGNRVSGNNSPSL 192  
  
QY 85 ESTGLEATFP-----KTTPLAQADPAGVGTPTGWCDCPLSDCTASAAGS-----STDDVE 134  
DB 193 SNGGFKPSRPPRPPTPRPASVNGSPSA-----TSEDGSGSTGSLPPTNTNINTE 248  
  
QY 135 LATE---FPATEMECELEGLLEERPALCLSP--QAPFPKLGWDELKPKCAQIYMRMQ 189  
DB 249 GATSGLIPLT-----ISGSGRPP---LNPVTOAPLPP--GWQRVDQHGRIYYVDHVE 298  
  
QY 190 EHTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPL---WDSKKQSFVGMILTIDF 246  
DB 299 KRTTWD-----RPELPFGWERRVDN--MGR1YYVDH 328  
  
QY 247 ILVLRHYRSPLVQIYEIQHK1ETWREIYVLCQCFPLVS---ISPNDSLF 294

DB 329 FTRTTTQWPTLESVRNYEQWLQ---RSQLOGAMQGFNRFIYGNODLF 375  
  
RESULT 14  
US-09-562-737-33  
; Sequence 33, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 830  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-33

Query Match 3.8%; Score 96; DB 4; Length 830;  
Best Local Similarity 24.5%; Pred. No. 0.32;  
Matches 46; Conservative 20; Mismatches 74; Indels 48; Gaps 9;  
  
QY 2 EPGLEHALRTPSWSLGGSEHQEMFLEQENSSSWPSPAVTSSSRIRGKRAKALRW 61  
DB 228 DPGIEHDLRSHSGGIEGRSSQELK-----SPG---SDSEDALGARLGRMISMI 274  
  
QY 62 RQKSVKEGEPGQGEGRSPRPAEAESTGLEATFPKTTPLAQADPAGVGTPTGWCDCPLSDC 121  
DB 275 SETELESNDGGSSSG-RSQHLTNSIE-EARSPASEPESEP--LHEPPRTAFLP--- 327  
  
QY 122 TASAAGSSTDVFLATEPATEWECLE-----EGLEER-----PALCL 161  
DB 328 -----VGQDDVN--SEYESSEWEPDUSEADADYPLLNLVYSIMISEGSPIDCQOCL 379  
  
QY 162 SPQAPFPK 169  
DB 380 SPAERLPE 387

RESULT 15  
US-07-582-945-2  
; Sequence 2, Application US/07582945  
; Patent No. 5369019  
; GENERAL INFORMATION:  
; APPLICANT: TAKKER FOGED, Niels  
; APPLICANT: PETERSON, Svend  
; TITLE OF INVENTION: A PASTEURELLA VACCINE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/582,945  
; FILING DATE: 19901012  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:

```
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30307/112 PLVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-582-945-2

Query Match      3.7%; Score 94.5; DB 1; Length 1285;
Best Local Similarity 19.9%; Pred. No. 1;
Matches 81; Conservative 58; Mismatches 109; Indels 159; Gaps 22;

QY 179 PGAQIYRMFOEHTCYDAMATSKLVI FDTMLEIKKAPFALVANGVRAAPLWDSKKQSFV 238
Db 162 PNEQIY----HSRVIADILYARS--VMD--EFKVFY-----EYWKYAQLYT 201
QY 239 GML-----TITD--FILVLRHYRS-----PLVQIYEIQH 267
Db 202 EMLSTFLAMALQOYTRQTLDGFLMVCNTYGNKEEVOITLLDIYGYPSTDIICIEQK 261
QY 268 KIETWKEI-YLQGCFKPLVSPNDSLFEAVYTLIKNRIHRLPV----- 310
Db 262 GLPTPKVILYIPGGTQPFVEFLNTDCLKQWIAHLKONKHVRFKHFSLKORQEGETFT 321
QY 311 -LD-----PVSQNVHLHILTHKRLKFLHFGSLPRPSFLYRTIIDLGI--- 353
Db 322 GIDKALQXIABESPENKANKYILYNPTH---LEWENLFNIMMKRTE--QRMLEDSDVQIR 376
QY 354 ---GTFRLAV-VLET-APILTALDIFV-----DPR 379
Db 377 SNSEATROYLSLLETISOLSAIDMLVPAGCIPINFALSATALGLSSDIVVNGDSYEKR 436
QY 380 -----VSALPVNRCQGVVGLYSRF--DVHLLAAQQTYNHLDMSVGEALR 422
Db 437 KYGIGSLVQSALFTGINLPVISETAELISLSPSRTEEDIPAFFEE-----QALA 486
QY 423 QRTLCLEGLV-SCOPHESLGEVIDRIAREQVHRLVY---DETQHLL 465
Db 487 QRFEIVEEBELHSISPDOPPREITD----ENLHKIRLVLRLANENQPLV 529
```

Search completed: June 6, 2003, 11:04:13  
Job time : 31.1173 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 11:02:59 ; Search time 40.2504 Seconds  
(without alignments)  
1254.259 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2538  
Sequence: 1 MEPGLEHALRTPSWSLGG.....LSDILQALVLPAGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues  
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published\_Applications\_AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2538	100.0	489	10	US-09-826-581-6
2	1017	40.1	344	10	US-09-925-297-461
3	892.5	35.2	1207	9	US-10-108-605-71
4	115	4.5	980	9	US-10-108-605-195
5	114	4.5	91	10	US-09-864-761-33979
6	105.5	4.2	1134	9	US-10-001-873-50
7	101	4.0	187	9	US-09-975-719-287
8	100.5	4.0	440	9	US-09-738-626-6006
9	100	3.9	1276	10	US-09-272-809-2
10	99.5	3.9	701	9	US-10-052-092-12
11	99.5	3.9	701	9	US-10-211-613-1
12	97.5	3.8	1139	9	US-09-971-490-5
13	97	3.8	739	9	US-10-097-534-10
14	96	3.8	830	9	US-10-211-962-33
15	96	3.8	1604	10	US-09-746-491-8
16	95.5	3.8	622	9	US-09-738-626-4919
17	94	3.7	830	9	US-10-211-962-37
18	93.5	3.7	419	9	US-10-257-963-12
19	93.5	3.7	419	10	US-09-175-254-3

20	93.5	3.7	2452	9	US-09-819-104A-5	Sequence 5, Appli
21	92.5	3.6	529	12	US-10-042-417-44	Sequence 44, Appl
22	92.5	3.6	1069	9	US-10-007-270-24	Sequence 24, Appl
23	92	3.6	595	9	US-10-052-092-31	Sequence 31, Appl
24	92	3.6	758	10	US-09-904-987-5	Sequence 5, Appli
25	91.5	3.6	660	10	US-09-853-033-8	Sequence 8, Appli
26	91	3.6	660	10	US-09-853-033-4	Sequence 4, Appli
27	90.5	3.6	909	9	US-10-226-296-4	Sequence 4, Appli
28	90.5	3.6	909	9	US-10-226-318-4	Sequence 4, Appli
29	90.5	3.6	1169	10	US-09-815-242-13693	Sequence 13693, A
30	90.5	3.6	1618	9	US-09-963-875-1	Sequence 1, Appli
31	90.5	3.6	1618	9	US-10-136-891-2	Sequence 2, Appli
32	90.5	3.6	1618	9	US-10-120-887-1	Sequence 1, Appli
33	90.5	3.6	2016	9	US-09-896-994-2	Sequence 2, Appli
34	90.5	3.6	2016	10	US-09-840-125-4	Sequence 4, Appli
35	90	3.5	581	9	US-09-298-523B-56	Sequence 56, Appl
36	90	3.5	595	10	US-09-853-033-2	Sequence 2, Appli
37	90	3.5	619	9	US-09-882-774-1	Sequence 1, Appli
38	90	3.5	807	9	US-10-044-692-5	Sequence 5, Appli
39	90	3.5	807	9	US-10-044-539-5	Sequence 5, Appli
40	90	3.5	2507	9	US-09-819-104A-2	Sequence 2, Appli
41	89.5	3.5	471	10	US-09-815-242-14028	Sequence 14028, A
42	89.5	3.5	660	10	US-09-853-033-6	Sequence 6, Appli
43	89.5	3.5	1399	9	US-09-388-221-4	Sequence 4, Appli
44	89.5	3.5	1424	9	US-09-388-221-12	Sequence 12, Appl
45	89.5	3.5	1429	9	US-10-028-392-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-09-826-581-6  
; Sequence 6, Application US/09826581  
; Patent No. US20020142310A1  
; GENERAL INFORMATION:  
; APPLICANT: Andersson, Leif  
; APPLICANT: Luthman, L. Holger  
; APPLICANT: Marklund, Stefan  
; TITLE OF INVENTION: VARIANTS OF THE HUMAN AMP-ACTIVATED PROTEIN KINASE GAMMA 3 SU.  
; FILE REFERENCE: 11145-007001  
; CURRENT APPLICATION NUMBER: US/09/826,581  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 60/195,665  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-581-6

Query Match	100.0%	Score	2538	DB	10	Length	489
Best Local Similarity	100.0%	Pred. No.	1.5e-209	Mismatches	0	Indels	0
Matches	489	Conservative	0	Gaps	0		
Qy	1	MEPGLHALRTPSWSLGGSEHQEMSFLEQENSSSWPSPAVTSSSERIRGKRAKALRW	60				
Db	1	MEPGLHALRTPSWSLGGSEHQEMSFLEQENSSSWPSPAVTSSSERIRGKRAKALRW	60				
Qy	61	TRQKSVEGEPGQGGPRSRPAEESTGLEATFPKTTPLAQADPAGVGPPTFGWDCPLPSD	120				
Db	61	TRQKSVEGEPGQGGPRSRPAEESTGLEATFPKTTPLAQADPAGVGPPTFGWDCPLPSD	120				
Qy	121	CTAAGSSTDDVELATEFPATEWCELEGLERPALCLSPQAPFFKLGWDELKRP	180				
Db	121	CTAAGSSTDDVELATEFPATEWCELEGLERPALCLSPQAPFFKLGWDELKRP	180				
Qy	181	AOIYRFRNQHTCYDAMATSSKLVIFDTMLEIKKAFYALVANGVRAAPLWDSKKQSVGM	240				
Db	181	AOIYRFRNQHTCYDAMATSSKLVIFDTMLEIKKAFYALVANGVRAAPLWDSKKQSVGM	240				

```
QY 241 LTTDTFLVHLHRYRSPVQIYEIOHKIETWREIYLOGCFKPLVSTSPNDSLFEAVYTL 300
Db 241 LTTDTFLVHLHRYRSPVQIYEIOHKIETWREIYLOGCFKPLVSTSPNDSLFEAVYTL 300
QY 301 IKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLA 360
Db 301 IKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLA 360
QY 361 VVLETAPLITALDIFVDRKVSALPVVNECGQVVGLYSRFDVIHLAAQOQTYNHLDMSVGEA 420
Db 361 VVLETAPLITALDIFVDRKVSALPVVNECGQVVGLYSRFDVIHLAAQOQTYNHLDMSVGEA 420
QY 421 LRORTLCLEGVLSOPHESIGEVDRIAREQVHRLVLDVDETOHLLGVVSLDILQALVLS 480
Db 421 LRORTLCLEGVLSOPHESIGEVDRIAREQVHRLVLDVDETOHLLGVVSLDILQALVLS 480
QY 481 PAGIDALGA 489
Db 481 PAGIDALGA 489

RESULT 2
US-09-925-297-461
; Sequence 461, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 461
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-461

Query Match 40.1%; Score 1017; DB 10; Length 344;
Best Local Similarity 64.8%; Pred. No. 3.6e-79;
Matches 193; Conservative 53; Mismatches 52; Indels 0; Gaps 0;

QY 183 IYKREMQHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKOSFVGMILT 242
Db 41 VYTSFKSHRCVDLIPSSKLVVFDVSLQVKKAFALVNTNGVRAAPLWDSKKOSFVGMILT 100
QY 243 ITDFILVHLHRYRSPVQIYEIOHKIETWREIYLOGCFKPLVSTSPNDSLFEAVYTLIK 302
Db 101 ITDFEINLHRYKSAQVILEBHKIETWREIYLDQSFKPLVCISPNASLFDVSSLLR 160
QY 303 NRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLLPSPFLYRTIQDLGIGTFRDLAV 362
Db 161 NKIHLPLVPDPSSGNTLYLTHKRLKFLKLFTEFPKPEFMSKSLLEELQIGTYANIAY 220
QY 363 LETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQOQTYNHLDMSVGEARL 422
Db 221 RTTTPVYVALGIFVQHRVSALPVVDEKGRVDIYSKFDVINLAAEKTYNNLVDVSKALQ 280
QY 423 QRTCLCLEGVLSOPHESIGEVDRIAREQVHRLVLDVDETOHLLGVVSLDILQALVLS 480
Db 281 HRSHYEGVLYKCYLHETLETINLVELAEVHRLVVVDENVVGVKIVSLSDILQALVLT 338

RESULT 3
US-10-108-605-71
; Sequence 71, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-71

Query Match 35.2%; Score 892.5; DB 9; Length 1207;
Best Local Similarity 40.6%; Pred. No. 1.1e-67;
Matches 202; Conservative 93; Mismatches 121; Indels 81; Gaps 10;

QY 7 HALRTPSWSSLGGSEHOEMSFLEQENSSSWSPAVTS-----SSER 48
Db 594 HSLGHSPSAAAAAAVQOTLVYERRRYSAAHSPSTSRHHSPVHQCMLMERRRSDYSVEQ 653
QY 49 IRGKRRAKALRWTRKSVSEEGEPGQGEPRSPRAAESTGLEATFPKTTPLAQADPAGVG 108
Db 654 IE-----QMKROOQLL-----PRS-----GRKISLHFYSNHAGRLVLGGG 688
QY 109 T-PPTGWDCLPDSCTASAGSSTDVDELATER-PATEAMECELEGLLEERPALCLSPQAP 166
Db 689 NRRTDVLCP-----LESIKEFEQISEANKRLLAELLKKE-----725
QY 167 FPKLGWDELKRGGAQIYMRFMQEHCTYDAMATSSKLVIFDTMLEIKKAFALVANGVRA 226
Db 726 -----DD-----SQIEVKFEFRHKCYDLIPTSAKLVVFDTLIVKKAFFALVYNGVRA 773
QY 227 APLWDSKKOSFYGMILTITDFILVHLHRYRSPVQIYEIOHKIETWREIYLOGCFKPLVS 286
Db 774 APLWSEKQOFVGMILTITDFIKLQNTYKSPNASMEQLEEHKLDTRRSV-LHNQVMPLVS 832
QY 287 ISPNDSLFEAVYTLIKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLLPSPFLYR 346
Db 833 IGPASLYDAIKLHRSRHLRPLVIDPATGCVLYILTHKRLRFLFLYINELPKPAYMOK 892
QY 347 TIODLGIGTFRDLAVVLETAPLITALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAA 406
Db 893 SLRELKIGTYNNIETADETTSIITALKKFFERRVSALPLVDSGRLVDIYAKFDVINLAA 952
QY 407 QOQYNNHLDMSVGEARLQRTCLCLEGVLSOPHESIGEVDRIAREQVHRLVLDVDETOHLLG 466
Db 953 EKYTNLDVLSLRKANEHNEFWEGVQKCNLDESLYTIMERIVRAEVHRLVVVDENRKRVTG 1012
QY 467 VVSLSDILQALVSPAG 483
Db 1013 IISLSDILLYLVLRPSG 1029

RESULT 4
US-10-108-605-195
; Sequence 195, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
```

```
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 195
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-195

Query Match          4.5%; Score 115; DB 9; Length 980;
Best Local Similarity 22.5%; Pred. No. 0.49;
Matches 103; Conservative 66; Mismatches 187; Indels 102; Gaps 22;

QY 85 ESTGLEATPKTTPLAQADPAGVGTPTGWDCLPDSCTAAGSSTDD--VELATEFPAT 142
DB 87 EKAGCDEIWEKTCQVGGDDPS-----VEITQDIVESSEDERFEDMSDTAPPI 133
QY 143 EAWCELEGL--LEERPALCLSPQAPPKLGWDDLRKFGAQLYMRMQE--HTCYDAMA 198
DB 134 ELPPCCLSLRLDISETIOSSLSTPLRKKLSMALE-----SESYIKLLNLFHVCELDLN 188
QY 199 TSKKLVIPTMLLEIKKAPFALVANGVRAAPLWDSKKQSPVGMLTITDFIL--VLHRYIRS 256
DB 189 TEGHLHLF----EIFKNFLLKNKALFEIMFADDTTFDVGCLVDPSVQPKKHQYLK 244
QY 257 PLVQIYEIQHK-----IETWREIYQGCFKPLVISPNDLSFEAVYTLI-KNRIHR 307
DB 245 QLAKFREAVPIKNDLLAKIHQTERVQYIQDIILPTPSVFVEDNMLNLTLSFFFNKVEI 304
QY 308 LPVL-----DPSGVNLHLTHKRLKLFHLHFGSLLP--RPSLYRTIQD 350
DB 305 VTMIQDERYLLDVAVLVTDPTGDAKRRDTVLFLKEFCNTAQNLPQGGKDSF-YKLTLC 363
QY 351 LGIGTFROLAVYLE-----TAPILTALDIFVDRVSPALPVNVECGVVGLYSRFDV-- 401
DB 364 LGILQALELTLVMDNKKTKSADILTAIVEF-----SPLVVRN--YTLNQANRREVER 415
QY 402 --THLAQOYTHLMDSVGEALRQRTCLGVLVS--CQPHSEL---GEVIDRIARQVHR 454
DB 416 MLLNIAIEQLNDSEPELGIAVQ-----LMGIWKILLEPENMLTEKGFLENFYKYSVQT 470
QY 455 LV-----LVP-----ETQHLAGVVSLSLILQALV 478
DB 471 LVAPVILNTIGDRPONEDYQTAQLLGIV--LDILSPCV 506

RESULT 5
US-09-864-761-33979
; Sequence 33979, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomlca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
```

```
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33979
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006966.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9
; OTHER INFORMATION: SKISSPROT HIT: P80385, EVALUE 2.00e-05
; OTHER INFORMATION: EST_HUMAN HIT: BE763984.1, EVALUE 2.00e-15
; OTHER INFORMATION: EST_HUMAN HIT: AW957020.1, EVALUE 1.00e-09
US-09-864-761-33979

Query Match          4.5%; Score 114; DB 10; Length 91;
Best Local Similarity 65.6%; Pred. No. 0.018;
Matches 21; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 452 VHRVLVDETHQLHLGVVSLSDILQALVSPAG 483
DB 1 VHRLLVWVNEADSIVGISLSLILQALILTPAG 32

RESULT 6
US-10-001-873-50
; Sequence 50, Application US/10001873
; Patent No. US20020160388A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and I
```

```

; FILE REFERENCE: DEX-0275
; CURRENT APPLICATION NUMBER: US/10/001,873
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,055
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252,496
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1134
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-873-50

Query Match          4.2%; Score 105.5; DB 9; Length 1134;
Best Local Similarity 23.3%; Pred. No. 4;
Matches 70; Conservative 46; Mismatches 113; Indels 71; Gaps 13;

Qy 38 PSPAVTSSSRIRKRRKALRWTRQKSVEEGEPGQGGPRSR-----PAAEST 87
      |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 68 PSPAGCGG-----GLLEQAALSATGSCAEPSECPDFVEGPEPRVDSPGRTEPCTAALDL 123
      |||| | | | | | | | | | | | | | | | | | | | | | | | |

Qy 88 GLEATFPKTTPLAADP--AGVGTPTGHDCLPSDCTASAGSSTDDELATEFFATEAM 145
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 124 GVQLT-PETLVEAKEEVEPVGVV--EAVPEEGLAQVAPSESQ-----PTLEMS 172
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 146 ECELEGLLEERPALCLSPAPFPFKGLGWDELKPKGAQIYMRFMQHTC-YDAMATSSKL- 203
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 173 DCDVAGGQCPS--LEPQAVPVLGSTCFLEEASSDOFPLSEDLPLAGMNLAAAEALP 230
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 204 -----VITDMLLEIKKAFVALVANGVRAAPLWDSKKQSFVGMLTITDFILVLRHY 254
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db 231 QARPLSPGAAGAQALEAESLVL-----EQSFLHGITL----- 267

Qy 255 RSLPVQIYEIQHKIETWRIYILOGCFKPLNSISPNDSLFEAVYTLIKNRIHLPLVDPV 314
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 ---LSEIAELELER-----RSQEMGAERALVARPSLESLLAAGSHMLREVLDC-FVVDPL 319

RESULT 7
US-09-975-719-287
; Sequence 287, Application US/09975719
; Publication No. US20030022349A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Rahne, Laurence G.
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: SEQUENCES AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/975,719
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-287

Query Match          4.0%; Score 101; DB 9; Length 187;
Best Local Similarity 25.0%; Pred. No. 0.68;
Matches 44; Conservative 14; Mismatches 46; Indels 72; Gaps 10;

Qy 38 PSPAVTSS-----SERIKGRKRAKALRWTRQ-----KSVEEGEPGQG 75
      || | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 PAPCTRSCPGRSGRGTSRSAGSDR-MGRANGARWTKRLPPRGRSLADAPAPCA 77
      || | | | | | | | | | | | | | | | | | | | | | | | |

Qy 76 EGRSRPAEST-GLEAT-----FPKTTPLAQAD-----PAGVGTPTPGMD---C 116
```

```

      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 78 AASRARPRASSTPGYRSTWKLPRRFRSPCCAATGSRSCRAGRPSAGCVPGGMSGTGC 137
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Qy 117 LPSD-----CTASAAGSSTDDELATEFFATEWECELEGLLEERPALCLSPQAPFP 168
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 138 APCSRCRPFCAAP-----PPSPAW-----PARSSAGSAPSP 167

RESULT 8
US-09-738-626-6006
; Sequence 6006, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6006
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6006

Query Match          4.0%; Score 100.5; DB 9; Length 440;
Best Local Similarity 19.8%; Pred. No. 2.6;
Matches 64; Conservative 62; Mismatches 130; Indels 67; Gaps 15;

Qy 180 GAQIYMRPMQHTCYDAMATSSKLVITDMLLEIKKAFVALVANGVRAAPLWDSKKQSFVG 239
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 47 GSASLLRVIDERALHNM-----LIMLRTLLDASAARFA---GAIAYNYMDSWANGIVL 97
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 240 MLTITDFI--LVLHRYRSPVLQIYEIQHKIETWRIYILOGCFK-----PLVS----- 286
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 98 AIWVWSLLTFVAVGVGEGRT----VGRKNPYSVMLRSAAVVLGSLAKILGPARGLIWIGNI 153
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 287 ISPNDSLFEAVYTLIKNRIHLPLVDPVSGNVHLHTHKRLKLLKFLHFGS-----LLPR 340
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 154 IAPGPGFRNGPYA---TEVELREMVDAIQEHGVIIEERMIQSFVFLASTTVRVQVMVPR 210
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 341 PSFLYRTIQDILGIGTFRDLAWLLETAPILTALDIFVDRRVSAIPWNE-CGVVWGLYSRF 399
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 211 PEMIW-----IESG-----KTAGQATAL--CVSRGHSRIPIGENVDDIIGIVLK 254
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 400 DVIHAAQQTYNHLD---MSVGEALRQRTLCLEGLVLSQPHESLGEVIDRIAREQVHRL 455
      ||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 255 DLV---QKTYATDGGKSVLDEVKREATFVPDS-----KSLDALQEMQEDHKHTA 303
      ||:| | | | | | | | | | | | | | | | | | | | | | | |

Qy 456 VLVDETOHLLGWSLSDLQALV 478
      |||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 304 ILVDEYGGVAGLISIEDILEEIV 326
      |||| | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-09-272-809-2
; Sequence 2, Application US/09272809
```



US-10-211-613-1

```
Query Match          3.9%; Score 99.5; DB 9; Length 701;
Best Local Similarity 17.2%; Pred. No. 6.4;
Matches 8; Conservative 66; Mismatches 126; Indels 197; Gaps 22;

QY 37 WSPAVTSSSRTGRKRAKALRWTKQSV---EPEPPGOGGPRSRPAESTGLEATF 93
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 WSPPLMIK-----RSKNSLSLSITADQMVSALLDAEPP-----ILYSEY 331
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 94 PKTTPLAQADPAGVCTPTTGWDCPLPSDCTASAGSSTDVDELATEFPATEWCELEGL 153
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 332 DTPRPSEASMMGLT-----NLADRELVHMI-----NNAKRVPGFV 368
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 154 EERPALCLSPQAPPFKLGWDELKPKCAQIYMKPMOEHTCYDAMATSSKLIVFDTMLEIK 213
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 D-----LTLDQVHLECAW-----LEILMIGLWRSMEHPVKLLFAPNLL--- 409
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 214 KAFFALVANGRAAPLWDSKKQSFVGMLTITDFILVLHRYRSPVLOIYEIOHKIETWR 273
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 -----LDRNQCKVEGWEIFDMLLATSSRFM----- 437
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 274 EIVLQG---CFKPLVSI SPNDSLFEAVYTLI-----KNRIHLPVLDPVSGNVLH 320
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 -MNLQGEFPVCLKSIILN-----SGVTFILSSTLKSLSEKQHIHR--VLDKITDTLIH 488
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 321 ILT-----HKRLKFLHIFCSLLPRPSFLYRTIQDLGIGTFRLAVLLETAPILTA 371
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 LMAKAGLTQQQHQRALQLLLLSHI-----RHMENQ----- 521
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 372 LDIFVDRVSALPVNVEGQVGLYSRFDVHIAAQOQTYNHLDMSVGEALRQRTICLEGV 431
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 522 -----KCV-VEGWEIFDML-LATSSRFMMNLQ-GEFF---VCLKSI 557
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 432 L-----SCQPIESLGEVIDRIAKEUHVRLV--LVDETQH 463
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 558 ILLNSGVYTFILSSTLKSLSEKDHIIHVLOKITDTLIHLMAKAGLTQQQH 607
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-971-490-5
: Sequence 5, Application US/09971490
: Publication No. US20030032091A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel A.
: APPLICANT: Curtis, Rory A. J.
: APPLICANT: Bandaru, Rajasekhar
: TITLE OF INVENTION: 48120, 23479 AND 46689, NOVEL HUMAN HYDROLASES
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: 10448-100001
: CURRENT APPLICATION NUMBER: US/09/971,490
: PRIOR FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 60/238,170
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 60/237,991
: PRIOR FILING DATE: 2000-10-05
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1139
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-971-490-5
```

```
Query Match          3.8%; Score 97.5; DB 9; Length 1139;
Best Local Similarity 19.8%; Pred. No. 19;
Matches 76; Conservative 59; Mismatches 129; Indels 119; Gaps 18;

QY 1 MEPLGHALRRTPSNLSGGSE--HQMSEFLQENSSWSPAVT-----SSSERIRG 51
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 LPVPLTFELSRFPFNQSLQPEKIHNLKEF-----PQIIYMDRYWYRSKELIRN 407
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 52 KRRA-----KALKRWTKQSVKEGEPGQGEGR-----KSRPAESTGLEATF 93
```

```
Db 408 KRECIRKLKEEIKLIQOKLERIVKYGSGPARFPLDMLKYVIVFASPKPASESCPPESDT 467
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 94 PKTTPLAQADPAGVCTPTTGWDCPLPSDCT---ASAGSSTDVDELATEFPATEWCELE 150
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 468 HMTLPSSV-----HCSVSDQTSKSTSTESSQDVE--STFSPSPEDSLPKSK 513
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 151 GLLEERPALCLSPQAPPFKLGWDE-----LRKPGAQIYMRPMQEBHTCYDAMATSSKL 203
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 PLTSSSSNEM-PSOPAPRIVTDEEINFVKTCLOQRSEIEQDIQDLKTCIASTTQTIEQ 572
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 204 VIFDTMLE-----IKKA-----FFALVANGVRAAPL-----WDS-KKQ 235
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 573 MYCDPLLRQVRYRLHAVLHVHGQANAGHWAYVYNQPSQMLKYNDISVTESSWEEVERD 632
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 236 SFVGMLTITDFILVLR-----YYSRPLV-----QIYEIE-----QHKI 269
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 633 SYGGLNVSAXCLMYINDKLPYNABEAPTESQDMSVEALSVELKHYTOEDNWRFEQEV 692
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 270 ETWREIYLQGCFKPLVLSISPNDS 292
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 693 EWEER--EQSCKIPQMESSINSS 713
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-10-097-534-10
: Sequence 10, Application US/10097534
: Publication No. US20030049607A1
: GENERAL INFORMATION:
: APPLICANT: GREENER, TSVIKA
: APPLICANT: MOSKOWITZ, HAIM
: APPLICANT: REISS, YUVAL
: APPLICANT: ALROY, IRIS
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
: TITLE OF INVENTION: MATURATION
: FILE REFERENCE: PLV-001.01
: CURRENT APPLICATION NUMBER: US/10/097,534
: CURRENT FILING DATE: 2002-03-12
: PRIOR APPLICATION NUMBER: 60/275,224
: PRIOR FILING DATE: 2001-03-12
: PRIOR APPLICATION NUMBER: 60/308,958
: PRIOR FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: 60/340,170
: PRIOR FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 71
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 739
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-097-534-10

Query Match          3.8%; Score 97; DB 9; Length 739;
Best Local Similarity 22.1%; Pred. No. 11;
Matches 64; Conservative 40; Mismatches 114; Indels 72; Gaps 15;

QY 29 LEQENSSWSPAVTSSSRIKGRRAKALRWTKQSVKEGEPGQGEGR-----SRPAA 84
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 LESEVVTNGETTCSESASQNDGSRKSDETRYSTNGS--DOPEDAGAGENRRVSGNNSPL 79
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 85 ESTGLEATFP---KTTFLAQADPAGVGTPTGWDCLPSCDTASAAGS-----STDVVE 134
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 SNGGFKPSRPPRPPPPPTPRPASVNGSPSA----TSESDGSSSTGCTPTNTNTNTSE 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 135 LATE---PPATEWCELEGLLEERFALCLSP--QAPFPKLGWDELKPKGQIYMRWQ 189
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 GATISGLIPLT-----ISGSGPRP---LNPVTQAPLPP-GWEQRVQDHGRVYVDVVE 185
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 190 EHTCYDAMATSSKLIVFDTMLEIKKAPFALVANGVRAAPL---WDSKKQSFVGMLTITDF 246
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 KRTTND-----RPELPPGWRERRVDN-MGRIVYVDH 215
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 247 ILVLHRYRSLVQIYEIOHKIETWREIYLQGCFKPLVS--ISPNDSLF 294
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:55:08 ; Search time 6.60595 Seconds  
(without alignments)  
827.023 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GKRRAKALRWTRQKSVEE.....PGQGGRSPRAAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	214	100.0	489	22	AA47679 PRKAG3. Homo sapi
2	210	98.1	464	22	AAE00223 Human AMPK gamma s
3	144	67.3	464	23	AAE22985 Pig PRKAG3 polymor
4	142	66.4	464	22	AAE00222 Pig AMPK gamma sub
5	142	66.4	464	23	AAE22984 Pig wild-type PRKA
6	142	66.4	464	23	AAE22987 Pig PRKAG3 polymor
7	142	66.4	464	23	AAE22988 Pig PRKAG3 polymor
8	142	66.4	514	22	AAE00224 Sus scrofa Prkag3
9	136	63.6	464	23	AAE22986 Pig PRKAG3 polymor
10	64.5	30.1	171	22	ABG03666 Novel human diago

11	63.5	29.7	61	21	AAG35967	Zea mays protein f
12	61.5	28.7	415	19	AAW71562	Human hepatocyte n
13	61.5	28.7	630	19	AAW71560	Human hepatocyte n
14	61.5	28.7	631	19	AAW71580	Human hepatocyte n
15	61.5	28.7	631	19	AAW71559	Human native hepat
16	61.5	28.7	866	22	ABB67322	Drosophila melanog
17	61.5	28.7	933	22	ABB67321	Drosophila melanog
18	61.5	28.7	958	22	ABB58729	Drosophila melanog
19	60.5	28.3	152	22	ABG26522	Novel human diago
20	60	28.0	73	22	AAU41694	Propionibacterium
21	60	28.0	187	20	AAV29189	Amino acid sequenc
22	60	28.0	270	20	AAU29328	Human secreted pro
23	60	28.0	270	22	AAU39054	Human secreted pro
24	60	28.0	270	23	ABB55763	Human polypeptide
25	59.5	27.8	1221	20	AAV05940	Thermophilus therm
26	59.5	27.8	1255	20	AAV05944	Thermophilus therm
27	59	27.6	230	22	ABG08090	Novel human diago
28	59	27.6	705	22	AAW39328	Human polypeptide
29	59	27.6	714	22	AAW41114	Human polypeptide
30	59	27.6	863	22	ABB64979	Drosophila melanog
31	58	27.1	144	22	ABG03531	Novel human diago
32	58	27.1	271	23	AAU96713	Human osteoclast a
33	58	27.1	282	23	AAU96712	Human osteoclast a
34	56.5	26.4	529	21	AAV83085	F-box protein FBP-
35	56.5	26.4	529	22	AAE48307	Human 2F26 protein
36	56	26.2	405	23	AAE14452	Human protein phos
37	56	26.2	447	23	AAU82751	Amino acid sequenc
38	56	26.2	449	22	ABG27792	Novel human diago
39	56	26.2	493	23	AAU75786	Human protein phos
40	55.5	25.9	1133	22	AAE04836	Human SGP018 phos
41	55.5	25.9	102	22	AAO10686	Human polypeptide
42	55.5	25.9	107	22	AAG92689	C glutamicum prote
43	55.5	25.9	345	21	AAV70506	Human BAG-1L prote
44	55.5	25.9	601	22	ABG23745	Novel human diago
45	55.5	25.9	678	22	AAW25601	Human protein sequ

ALIGNMENTS

RESULT 1	
AA47679	
ID	AA47679 standard; Protein: 489 AA.
AC	AA47679;
XX	
DT	21-JAN-2002 (first entry)
DE	PRKAG3.
XX	
KW	Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW	metabolic disease; diabetes; obesity; substitution; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 71 /note= "Possible variation point P71A"
FT	Misc-difference 340 /note= "Possible variation point R340W"
XX	
PN	WO200177305-A2.
XX	
PD	18-OCT-2001.
XX	
PF	06-APR-2001; 2001WO-SE00765.
XX	
PR	07-APR-2000; 2000US-195665P.
XX	
PA	(AREX-) AREXIS AB.
XX	
PI	Andersson L, Luthman H, Marklund S;
XX	

DR WPI: 2001-657170/75.  
 XX N-PSDB; AAH43685.  
 XX  
 PT New variants of human AMP-activated protein kinase gamma3 subunit  
 PT associated with a metabolic disease e.g. diabetes or obesity and method  
 PT for determining a risk estimate of diseases in subject by detecting the  
 PT variant -  
 XX  
 PS Disclosure: Fig 5; 25pp; English.  
 XX  
 CC This sequence is encoded by the full length cDNA encoding the human  
 CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting  
 CC the presence of the PRKAG3 DNA, or a variant, is useful in determining  
 CC a risk estimate of a metabolic disease, such as diabetes or obesity,  
 CC in a subject. The variation may occur in exons 3, 4 or 10. In exon  
 CC 3 variation may be a substitution of a G for a C at nucleotide 320,  
 CC resulting in the amino acid substitution P71A; in exon 4 variation may  
 CC be a substitution of a T for a C at nucleotide 550; and in exon 10  
 CC variation may be a substitution of a T for a C at nucleotide 1037,  
 CC resulting in the amino acid substitution R340W. There may also be  
 CC nucleotide variation in intron 6. The numbering of these  
 CC variations is based on the full length cDNA, rather than on  
 CC position 1 of the open reading frame.  
 XX  
 SQ Sequence 489 AA;  
 Query Match 100.0%; Score 214; DB 22; Length 489;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-19;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GKRRAKALRWTRQKSVGEPPGQGGPRSPRAESTGLEA 41  
 |||||||  
 Db 51 GKRRAKALRWTRQKSVGEPPGQGGPRSPRAESTGLEA 91  
 RESULT 2  
 AAEE00223  
 ID AAEE00223 standard; Protein: 464 AA.  
 XX  
 AC AAEE00223;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
 XX  
 KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cadiant; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 172..225  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Misc-difference 200  
 FT /note= "RN- mutation site"  
 FT Domain 253..307  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 329..382  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 400..453  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 XX  
 WC200120003-A2.  
 XX  
 XX 22-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-EP09896.

XX  
 PR 10-SEP-1999; 99EP-0402236.  
 PR 18-MAY-2000; 2000EP-0401388.  
 XX  
 PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX  
 DR WPI: 2001-244810/25.  
 DR N-PSDB; AAD03320.  
 XX  
 PT New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 CC Claim 5; Fig 3; 71pp; English.  
 CC  
 CC The present sequence is human adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform, of  
 CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of  
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
 CC useful as therapeutic for treating carbohydrate metabolism disorders such  
 CC as diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 98.1%; Score 210; DB 22; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 1.9e-18;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GKRRAKALRWTRQKSVGEPPGQGGPRSPRAESTGLEA 41  
 |||||||  
 Db 26 GKRRAKALRWTRQKSVGEPPGQGGPRSPRAESTGLEA 66  
 RESULT 3  
 AAEE22985  
 ID AAEE22985 standard; Protein: 464 AA.  
 XX  
 AC AAEE22985;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig PRKAG3 polymorphic variant (PRKAG3-30).  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 30  
 FT /note= "Wild type Asn is substituted with Thr due

FT to single nucleotide polymorphism (SNP)"

XX

PN W0200220850-A2.

XX

XX 14-MAR-2002.

XX

XX 10-SEP-2001; 2001WO-US28283.

XX

XX 08-SEP-2000; 2000US-231045P.

PR 08-JAN-2001; 2001US-260239P.

PR 18-JUN-2001; 2001US-299111P.

XX

PA (LOWA ) UNIV IOWA STATE RES FOUND INC.

XX

XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;

PI

XX WPI; 2002-393850/42.

DR N-PSDB; AAD36457.

XX

XX Screening animals to determine those likely to produce larger litters

FT and improved meat quality traits involves assaying for the presence of

PT polymorphisms in the AMP activated protein kinase regulatory gamma

PT subunit gene .

XX

XX Disclosure; Page 91-93; 109pp; English.

PS

XX

CC The invention relates to a method for screening animals to determine

CC those more likely to produce large litters and improved meat quality

CC traits. The method involves assaying for the presence of a genotype

CC in the sample of genetic material obtained from animal. The genotype

CC is characterised by polymorphism(s) in the AMP activated protein

CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used

CC for screening animals e.g., pigs to determine those most likely to

CC exhibit improved meat quality traits and to produce larger litters.

CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-30).

XX

SQ Sequence 464 AA;

Query Match 67.3%; Score 144; DB 23; Length 464;

Best Local Similarity 68.3%; Pred. No. 4.7e-10;

Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 1 GKRRKALRWTRQKSVVEEGPPQGEGPRSPRAESTGLEA 41

Db | : || ||||| : ||||| ||| ||||| ||||| ||

26 GDQGTAKSRWTRQEDVEEGPPGPRGPGQSRPVAESTGQEA 66

RESULT 4

AAE00222

ID AAE00222 standard; Protein: 464 AA.

XX

XX AAE00222;

XX

XX 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.

XX

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;

KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;

KW cystathione beta synthase; CBS; cardiac; gene therapy; RN locus;

XX chromosome 15.

XX

OS Sus scrofa.

XX

XX

EH Key Location/Qualifiers

FT Domain 172..225

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Misc-difference 200

FT /note= "RN- mutation site"

FT Domain 253..307

FT /label= CBS

FT

FT Domain

FT /note= "Cystathione beta synthase domain"

FT 329..382

FT /label= CBS

FT

FT /note= "Cystathione beta synthase domain"

FT 400..453

FT /label= CBS

FT

FT /note= "Cystathione beta synthase domain"

XX

XX W0200120003-A2.

PN

XX 22-MAR-2001.

PD

XX

PF 11-SEP-2000; 2000WO-EP09896.

XX

XX 10-SEP-1999; 99EP-0402236.

PR 18-MAY-2000; 2000EP-0401388.

XX

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA (ANDE/) ANDERSSON L.

PA (LOOF/) LOOFT C.

PA (KALM/) KALM E.

XX

PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;

XX

DR WPI; 2001-244810/25.

DR N-PSDB; AAD03319.

XX

XX New variants of the gamma subunit of vertebrate adenosine

PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and

PT myopathy -

XX

PS Claim 5; Fig 3; 71pp; English.

XX

CC The present sequence is pig adenosine monophosphate

CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,

CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome

CC 15. Mutation in Prkag3 results in an altered regulation of carbohydrate

CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as

CC therapeutic for treating carbohydrate metabolism disorders such as

CC diabetes, obesity, and disorders associated with muscle metabolism

CC such as myopathy and cardiovascular diseases, to modulate AMPK

CC activity, and for restoring a normal AMPK function. PRKAG3 sequence

CC and its functionally altered mutants are useful for the diagnostic

CC evaluation, genetic testing and prognosis of a metabolic disorder.

CC preferably a carbohydrate metabolism disorder. Primers that can detect

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3 are

CC useful for detecting a dysfunction of carbohydrate metabolism resulting

CC from the expression of a functionally altered allele of PRKAG3.

CC Transgenic animal and host cell transformed with PRKAG3 or a

CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

CC screening compounds able to modulate AMPK activity. Nucleic acid

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or

CC in a sequence encoding the first cystathione beta synthase (CBS) domain

CC of PRKAG3 and is useful in gene therapy.

XX

SQ Sequence 464 AA;

Query Match 66.4%; Score 142; DB 22; Length 464;

Best Local Similarity 68.3%; Pred. No. 8.4e-10;

Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 1 GKRRKALRWTRQKSVVEEGPPQGEGPRSPRAESTGLEA 41

Db | : || ||||| : ||||| ||| ||||| ||||| ||

26 GDQGNKASRWTRQEDVEEGPPGPRGPGQSRPVAESTGQEA 66

RESULT 5

AAE22984

ID AAE22984 standard; Protein: 464 AA.

XX

XX AAE22984;

```

XX 09-AUG-2002 (first entry)
XX Pig wild-type PRKAG3 protein.
XX
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
KW screening; meat quality; single nucleotide polymorphism; SNP; pig.
XX
XX Sus scrofa.
XX
XX Key Location/Qualifiers
FH /note= "Wild type Asn is replaced with Thr during
FT single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 52
FT /note= "Wild type Gly is replaced with Ser during
FT single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 199
FT /note= "Wild type Val is replaced with Ile during
FT single nucleotide polymorphism (SNP)"
FT
FT Misc-difference 200
FT /note= "Wild type Arg is replaced with Gln during
FT single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
XX
XX 14-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US28283.
XX
XX 08-SEP-2000; 2000US-231045P.
PR 08-JAN-2001; 2001US-260239P.
PR 18-JUN-2001; 2001US-299111P.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
XX WPI; 2002-393850/42.
DR N-PSDB; AAD36456.
XX
XX Screening animals to determine those likely to produce larger litters
PT and improved meat quality traits involves assaying for the presence of
PT polymorphisms in the AMP activated protein kinase regulatory gamma
PT subunit gene -
XX
XX Claim 2; Fig 1; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype
CC in the sample of genetic material obtained from animal. The genotype
CC is characterised by polymorphism(s) in the AMP activated protein
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used
CC for screening animals e.g., pigs to determine those most likely to
CC exhibit improved meat quality traits and to produce larger litters.
CC The present sequence is pig wild-type PRKAG3 protein.
XX
XX Sequence 464 AA;
SQ
Query Match 66.4%; Score 142; DB 23; Length 464;
Best Local Similarity 66.3%; Pred. No. 8.4e-10;
Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GKRRKALRWTRQKSVEEGPPGQGGPRSRPAAESTGLEA 41
| : || |||||: |||| ||| ||||: |||| |||| ||
Db 26 GDQGNKASRWTRQEDVEEGPPGPRGQSRPVAESTGQEA 66

RESULT 6
AAE22987
ID AAE22987 standard; Protein: 464 AA.
XX
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;

```

```

AC AAE22987;
XX
XX 09-AUG-2002 (first entry)
XX
XX Pig PRKAG3 polymorphic variant (PRKAG3-199).
XX
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
KW screening; meat quality; single nucleotide polymorphism; SNP; pig;
XX variant.
XX
XX Sus scrofa.
XX
XX Key Location/Qualifiers
FH Misc-difference 199
FT /note= "Wild type Val is substituted with Ile due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
XX
XX 14-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US28283.
XX
XX 08-SEP-2000; 2000US-231045P.
PR 08-JAN-2001; 2001US-260239P.
PR 18-JUN-2001; 2001US-299111P.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
XX WPI; 2002-393850/42.
DR N-PSDB; AAD36459.
XX
XX Screening animals to determine those likely to produce larger litters
PT and improved meat quality traits involves assaying for the presence of
PT polymorphisms in the AMP activated protein kinase regulatory gamma
PT subunit gene -
XX
XX Disclosure; Page 100-102; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype
CC in the sample of genetic material obtained from animal. The genotype
CC is characterised by polymorphism(s) in the AMP activated protein
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used
CC for screening animals e.g., pigs to determine those most likely to
CC exhibit improved meat quality traits and to produce larger litters.
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-199).
XX
XX Sequence 464 AA;
SQ
Query Match 66.4%; Score 142; DB 23; Length 464;
Best Local Similarity 66.3%; Pred. No. 8.4e-10;
Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GKRRKALRWTRQKSVEEGPPGQGGPRSRPAAESTGLEA 41
| : || |||||: |||| ||| ||||: |||| |||| ||
Db 26 GDQGNKASRWTRQEDVEEGPPGPRGQSRPVAESTGQEA 66

RESULT 7
AAE22988
ID AAE22988 standard; Protein: 464 AA.
XX
XX AAE22988;
XX
XX 09-AUG-2002 (first entry)
XX
XX Pig PRKAG3 polymorphic variant (PRKAG3-200).
XX
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;

```

```

KW screening; meat quality; single nucleotide polymorphism; SNP; pig;
KW variant.
XX
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Misc-difference 200
FT /note= "Wild type Arg is substituted with Gln due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
XX
XX 14-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US28283.
XX
XX 08-SEP-2000; 2000US-231045P.
XX 08-JAN-2001; 2001US-260239P.
XX 18-JUN-2001; 2001US-299111P.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Rothschild WF, Ciobanu DC, Malek M, Plastow G;
XX
XX WPI; 2002-393850/42.
XX N-PSDB; AAD36460.
XX
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene -
XX
XX Disclosure; Page 105-107; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).
XX
XX Sequence 464 AA;
XX
Query Match 66.4%; Score 142; DB 23; Length 464;
Best Local Similarity 68.3%; Pred. No. 8.4e-10;
Matches 2b; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GKRRKALRWTRQKSVGEPPGCGGPRSRPAAESTGLEA 41
   | : || ||||| : ||||| ||| ||||| ||||| ||
Db 26 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 66

RESULT 8
AAE00224
ID AAE00224 standard; Protein; 514 AA.
XX
XX AAE00224;
XX
XX 13-JUN-2001 (first entry)
XX
XX Sus scrofa Prkag3 splice variant.
XX
XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
XX PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
XX genetic testing; carbohydrate metabolism disorder; skeletal muscle;
XX cystathione beta synthase; CBS; cardiatic; gene therapy.
XX
XX Sus scrofa.
XX
XX WO200120003-A2.
XX

```

```

XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-EP09896.
XX
XX 10-SEP-1999; 99EP-0402236.
XX 18-MAY-2000; 2000EP-0401388.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFT C.
XX (KALM/) KALM E.
XX
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX Iannuccelli N, Geillin J, Le Roy P, Chardon P;
XX
XX WPI; 2001-244810/25.
XX N-PSDB; AAD03321.
XX
XX New variants of the gamma subunit of vertebrate adenosine
XX monophosphate-activated kinase for diagnosis or treatment of disorders
XX associated with energy metabolism such as diabetes, obesity, and
XX myopathy -
XX
XX Claim 5; Page 70-71; 71pp; English.
XX
XX The present sequence is pig adenosine monophosphate (AMP)-activated
XX kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
XX variant. Prkag3 gene is located in the RN locus of chromosome 15.
XX Mutation in Prkag3 results in an altered regulation of carbohydrate
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX therapeutic for treating carbohydrate metabolism disorders such as
XX diabetes, obesity, and disorders associated with muscle metabolism
XX such as myopathy and cardiovascular diseases, to modulate AMPK
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX and its functionally altered mutants are useful for the diagnostic
XX evaluation, genetic testing and prognosis of a metabolic disorder,
XX preferably a carbohydrate metabolism disorder. Primers that can detect
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX useful for detecting a dysfunction of carbohydrate metabolism resulting
XX from the expression of a functionally altered allele of PRKAG3.
XX Transgenic animal and host cell transformed with PRKAG3 or a
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX screening compounds able to modulate AMPK activity. Nucleic acid
XX encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
XX in a sequence encoding the first cystathione beta synthase (CBS) domain
XX of PRKAG3 and is useful in gene therapy.
XX
XX Sequence 514 AA;
XX
Query Match 66.4%; Score 142; DB 22; Length 514;
Best Local Similarity 68.3%; Pred. No. 9.3e-10;
Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GKRRKALRWTRQKSVGEPPGCGGPRSRPAAESTGLEA 41
   | : || ||||| : ||||| ||| ||||| ||||| ||
Db 76 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 116

RESULT 9
AAE22986
ID AAE22986 standard; Protein; 464 AA.
XX
XX AAE22986;
XX
XX 09-AUG-2002 (first entry)
XX
XX Pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
XX screening; meat quality; single nucleotide polymorphism; SNP; pig;
XX variant.
XX

```

```
OS Sus scrofa.
XX Key Location/Qualifiers
FH Misc-difference 52
FT /note= "Wild type Gly is substituted with Ser due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
XX
XX 14-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US28283.
XX
XX 08-SEP-2000; 2000US-231045P.
PR 08-JAN-2001; 2001US-260239P.
PR 18-JUN-2001; 2001US-299111P.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
XX WPI; 2002-393850/42.
XX N-PSDB; AAO36458.
XX
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene -
XX
XX Claim 36; Page 96-97; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX Sequence 464 AA:
XX
XX Query Match 63.68; Score 136; DB 23; Length 464;
XX Best Local Similarity 65.98; Pred. No. 4.8e-09;
XX Matches 27; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
XX
XX QY 1 GRRKALRWTRQKSVEEGPPGQGGPRSRPAESTGLEA 41
XX : | | | | | | | | | | | | | | | | | | | |
XX 26 GDQGNKASRWTRQEDVEEGPGPGPRSPQSRPVAESTQGEA 66
XX
XX RESULT 10
XX ABG03666
XX ID ABG03666 standard; Protein; 171 AA.
XX
XX AC ABG03666;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #3657.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX
```

```
PR 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS67853.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 34025; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 171 AA:
XX
XX Query Match 30.1%; Score 64.5; DB 22; Length 171;
XX Best Local Similarity 37.2%; Pred. No. 2;
XX Matches 16; Conservative 5; Mismatches 17; Indels 5; Gaps 1;
XX
XX QY 3 RRAKALRWTRQKSVEEGPPGQGGEG-----PRSRPAESTGLE 40
XX : | | | | | | | | | | | | | | | | | | | |
XX 119 RVARGPRSTRSSAVDAGPPPSASPGFPXSSSQQRSPSEKTSGE 161
XX
XX RESULT 11
XX AAG35967
XX ID AAG35967 standard; Protein; 61 AA.
XX
XX AC AAG35967;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Zea mays protein fragment SEQ ID NO: 44011.
XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX OS Zea mays subsp. mays.
XX
XX PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX
XX PF 25-FEB-1999; 99US-0121825.
XX
```

PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.







```
XX
PN WO9811254-A1.
XX
PD 19-MAR-1998.
XX
PF 10-SEP-1997; 97WO-US16037.
XX
PR 30-OCT-1996; 96US-0029679.
PR 10-SEP-1996; 96US-0025719.
PR 02-OCT-1996; 96US-0028056.
XX
PA (ARCH-) ARCH DEV CORP.
XX
XX Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S;
PI Oda N, Yamagata K;
XX
DR WPI; 1998-271667/24..
DR N-PSDB; AAV52625.
XX
XX Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and
PT 1-beta - useful for detecting susceptibility for non-insulin
PT dependent diabetes, especially maturity-onset diabetes of the young
XX
PS Claim 14; Fig 11; 363pp; English.
XX
CC This is the amino acid sequence of human hepatocyte nuclear factor 1
CC alpha (HNF-1 alpha), a homeodomain-containing sequence-specific DNA
CC binding protein. A cDNA sequence (see AAV52624) encoding HNF-1 alpha
CC is provided. Mutations in this transcription factor are indicative
CC of a propensity to diabetes mellitus. The HNF-1 alpha gene is
CC located on human chromosome 12q, which is the location site of the
CC MODY3 (maturity onset diabetes of the young) locus. The invention
CC concerns the identification of genes responsible for non-insulin
CC dependent diabetes mellitus (NIDDM) for use in diagnostics and
CC therapeutics. It demonstrates that the MODY3 locus is the HNF-1
CC alpha gene, the MODY4 locus is the HNF-1 beta gene (see AAV52730) and
CC the MODY1 locus is the HNF-4 alpha gene (see AAV52687). Analysis of
CC mutations in these HNF genes can be diagnostic for diabetes. The
CC invention also contemplates methods of screening for modulators of
CC HNF function utilising HNF nucleic acids or polypeptides, the
CC modulators being useful for treating diabetes by modulating HNF
CC function in an animal.
XX
SQ Sequence 631 AA;
Query Match 28.7%; Score 61.5; DB 19; Length 631;
Best Local Similarity 39.5%; Pred. No. 20;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;
Qy 2 KRAKALWTRQKSVEEGPPQGGGPRSRPAESTGL 39
Db :|:| | | | | | | | | | | | | |
271 RKKEAFRHLAMDYSGPPPGPGCP-ALPAHSSPGL 307
Search completed: June 6, 2003, 11:01:02
Job time : 8.60596 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:58:29 ; Search time 2.65674 Seconds  
(without alignments)  
1483.588 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91  
Perfect score: 214  
Sequence: 1 GKRRAKALRWTRQKSVKEGE.....PQCGCPRSRPAESTGLEA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	29.0	493	2 T05105	hypothetical prote
2	61.5	28.7	494	2 S39607	transcription fact
3	61.5	28.7	542	2 S39608	transcription fact
4	61.5	28.7	628	1 A39262	transcription fact
5	61.5	28.7	628	1 A33313	transcription fact
6	61.5	28.7	631	1 A36749	transcription fact
7	59	27.6	705	2 A35363	synapsin I splice
8	57.5	26.9	565	2 AF3639	pseudouridylylate sy
9	57	26.5	149	1 R5MXE	ribosomal protein
10	56	26.2	435	2 H87516	D-alanyl-D-alanine
11	56	26.2	969	2 A75634	McrB-related prote
12	55.5	25.9	673	2 S36336	probable transcrip
13	55.5	25.9	943	2 T34847	probable transcrip
14	55	25.7	691	2 A25704	synapsin I - rat
15	55	25.7	704	2 A30411	synapsin Ia - rat
16	55	25.7	1085	2 F96712	hypothetical prote
17	55	25.7	1380	2 S34307	myosin heavy chain
18	54.5	25.5	858	2 S88227	ubiquitin thiolest
19	54.5	25.5	1240	2 S21086	anion exchange pro
20	54	25.2	171	2 S55461	mucin - human (fra
21	54	25.2	396	2 D69378	3-ketoacyl-CoA thi
22	54	25.2	701	2 S06475	phenylalanine ammo
23	53.5	25.0	1087	2 T31100	probable potassium
24	53	24.8	203	2 S76108	hypothetical prote
25	52.5	24.5	206	2 B48441	antigen (C-termina
26	52.5	24.5	275	1 JG1113	interleukin-2 rece
27	52.5	24.5	277	2 T51975	C 3.4.25.1 proteas
28	52.5	24.5	303	2 E86990	protein C36C5.13 [
29	52.5	24.5	323	2 S16318	homeotic protein H

30	52.5	24.5	356	1 WJHU2H	homeotic protein H
31	52.5	24.5	660	1 Q0SE3	BHLF1 protein - hu
32	52	24.3	78	2 C95995	hypothetical prote
33	52	24.3	369	1 C70666	probable membrane-
34	52	24.3	517	2 T28630	Y4cC protein - Rhi
35	52	24.3	612	2 S53714	probable dinitrifi
36	52	24.3	751	2 T46517	hypothetical prote
37	52	24.3	1234	2 A34911	band 3-related pro
38	52	24.3	1237	2 A31789	band 3-related pro
39	52	24.3	1237	2 A56764	band 3-related pro
40	52	24.3	1280	2 T00365	hypothetical prote
41	51.5	24.1	200	2 S14277	NADH2 dehydrogenas
42	51.5	24.1	259	2 T34536	hypothetical prote
43	51.5	24.1	363	1 OZZQAK	circumsporozoite p
44	51.5	24.1	1359	2 T34036	hypothetical prote
45	51.5	24.1	1530	1 S01393	DNA-directed RNA p

ALIGNMENTS

RESULT 1

T05105

hypothetical protein F28M20.160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Mar-2002

C:Accession: T05105

R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mew

submitted to the Protein Sequence Database, November 1998

A:Reference number: 215398

A:Accession: T05105

A:Molecule type: DNA

A:Residues: 1-493 <BEV>

A:CROSS-references: EMBL:AL031004

C:Experimental source: cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:introns: 26/3; 105/1; 146/3; 296/3; 406/3

A:Note: F28M20.160

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 29.0%; Score 62; DB 2; Length 493;

Best Local Similarity 35.8%; Pred. No. 9.1;

Matches 16; Conservative 7; Mismatches 8; Indels 14; Gaps 2;

Qy 3 RRAKALRWTRQKSVKEGEPPGCGCPRSRPAESTG 38

Db 238 KRARVQWQSQESRSKHDVREKTAEEGEP-----SRNRITSNKSTG 277

RESULT 2

S39607

transcription factor HNF-1A, splice form C - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 29-Aug-1997

C:Accession: S39607; S39752

R:Bach, I.; Yaniv, M.

EMBO J. 12, 4229-4242, 1993

A:Title: More potent transcriptional activators or a transdominant inhibitor of the

A:Reference number: S39607; PMID:94038905; PMID:7900999

A:Accession: S39607

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-494 <BAC>

A:CROSS-references: EMBL:X71347

R:Bach, I.

submitted to the EMBL Data Library, April 1993

A:Reference number: S39752

A:Accession: S39752

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 118-494 <BA2>

A:CROSS-references: EMBL:X71347





Matches 15; Conservative 5; Mismatches 7; Indels 14; Gaps 2;

QY 15 SVREGPPGGEGPR---SRPAEST-----GLEA 41  
::||| i::||| |||

Db 320 TLEAGLDGRCDFRQAOCFARQHTLLTALNAPFLAGLEA 360  
::||| i::||| |||

RESULT 12  
S36336  
probable transcription factor PLZF - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 05-Nov-1999  
C:Accession: S36336; S31989  
R:Chen, Z.; Brand, N.J.; Chen, A.; Chen, S.J.; Tong, J.H.; Wang, Z.Y.; Waxman, S.; Z.  
EMBO J. 12, 1161-1167, 1993  
A:Title: Fusion between a novel Krueppel-like zinc finger gene and the retinoic acid  
A:Reference number: S36336; PMID:93209216; PMID:8384553  
A:Accession: S36336  
A:Molecule type: mRNA  
A:Residues: 1-673 <CHE>  
A:Cross-references: EMBL:Z19002; NID:g38517; PIDN:CAA79489.1; PID:g38518  
C:Genetics:  
A:Gene: PLZF  
C:Superfamily: POZ domain homology  
C:Keywords: zinc finger  
F:20-118/Domain: POZ domain homology <POZ>

Query Match 25.9%; Score 55.5; DB 2; Length 673;  
Best Local Similarity 37.9%; Pred. No. 71;  
Matches 11; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 5 AKALWTRQKSVSEEGPQG-GEGRSRP 32  
i::||| i::|||

Db 292 ARELHYGREBSAQVPPPAEAGAPTGRP 320  
i::||| i::|||

RESULT 13  
T34847  
probable transcription regulator - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34847  
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
submitted to the EMBL Data Library, February 1999  
A:Reference number: Z21559  
A:Accession: T34847  
A>Status: preliminary; translated from GB/EMBL/DDBB  
A:Molecule type: DNA  
A:Residues: 1-943 <OLI>  
A:Cross-references: EMBL:AL035478; PIDN:CAB36601.1; GSPDB:GN00070; SCOEDB:SC2G5.14c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC2G5.14c

Query Match 25.9%; Score 55.5; DB 2; Length 943;  
Best Local Similarity 40.5%; Pred. No. 97;  
Matches 15; Conservative 3; Mismatches 16; Indels 3; Gaps 1;

QY 3 BRKALEPWTKSVSEEGPPGCGEGRSRPAESTGL 39  
||:||| i::|||

Db 847 RRCRALFWLR---VEAAVAVRASGPAPVPAEPDGL 880  
||:||| i::|||

RESULT 14  
A25704  
synapsin I - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 21-May-1998 #sequence\_revision 21-May-1988 #text\_change 05-Nov-1999  
C:Accession: A25704  
R:McCaffery, C.A.; deGennaro, L.J.  
EMBO J. 5, 3167-3173, 1986  
A:Title: Determination and analysis of the primary structure of the nerve terminal s  
A:Reference number: A25704; PMID:87133474; PMID:3028773





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 1.50788 Seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: us-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GKRRKALRWTRQKSVSEGE.....PQGGGPRSPRAESTGLEA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	210	98.1	464	1 AAK1_HUMAN	Q9ug19 homo sapien
2	142	66.4	464	1 AAK1_PIG	Q9m9p4 sus scrofa
3	61.5	28.7	628	1 HNFA_MOUSE	P22361 mus musculus
4	61.5	28.7	628	1 HNFA_RAT	P15257 rattus norv
5	61.5	28.7	631	1 HNFA_HUMAN	P20823 homo sapien
6	61.5	28.7	958	1 HIG_DROME	Q09101 drosophila
7	59	27.6	275	1 CAB4_HUMAN	P57796 homo sapien
8	59	27.6	705	1 SYN1_HUMAN	P17600 homo sapien
9	57	26.6	149	1 RL19_METVA	P14024 methanococc
10	55.5	25.9	569	1 U171_HUMAN	Q12980 homo sapien
11	55.5	25.9	673	1 Z145_HUMAN	Q05516 homo sapien
12	55	25.7	270	1 SLBP_HUMAN	Q14493 homo sapien
13	55	25.7	670	1 SYN1_MOUSE	O88935 mus musculu
14	55	25.7	704	1 SYN1_RAT	P09951 rattus norv
15	55	25.7	880	1 G0N4_THEFU	P26221 thermomonos
16	55	25.7	1980	1 MY9B_RAT	Q63358 rattus norv
17	54.5	25.5	299	1 CRX_HUMAN	O43186 homo sapien
18	54.5	25.5	858	1 UBP5_HUMAN	P45974 homo sapien
19	54.5	25.5	858	1 UBP5_MOUSE	P56399 mus musculu
20	54	25.2	701	1 PAL1_ORYSA	P14717 oryza sativ
21	54	25.2	1739	1 CHD2_HUMAN	O14647 homo sapien
22	53	24.8	461	1 Y514_HUMAN	O60269 homo sapien
23	53	24.8	830	1 SREC_HUMAN	Q14162 homo sapien
24	52.5	24.5	275	1 IL2A_SHEEP	P26898 ovis aries
25	52.5	24.5	277	1 PS12_ARATH	Q23712 arabidopsis
26	52.5	24.5	299	1 CRX_BOVIN	Q9xsk0 bos taurus
27	52.5	24.5	299	1 CRX_MOUSE	O54751 mus musculu
28	52.5	24.5	323	1 HXDB_MOUSE	P23813 mus musculu
29	52.5	24.5	356	1 HXB2_HUMAN	P14652 homo sapien
30	52.5	24.5	660	1 YHL1_EBV	P03181 Epstein-bar
31	52	24.3	369	1 MDCQ_MYCTU	P95155 mycobacteri
32	52	24.3	415	1 SYN1_CANFA	O62732 canis famil
33	52	24.3	517	1 YACC_RHISN	P55385 thizobium s

RESULT 1									
AAKI_HUMAN	1	528	1	PPGA_HUMAN	1	528	1	Q96134	homo sapien
AC	09UG19; Q9NRL1;	24.3	612	1 Y525_PSEAE	1	612	1	Q51484	pseudomonas
DT	16-OCT-2001 (Rel. 40, Created)	24.3	1234	1 B3A2_RAT	1	1234	1	P23347	rattus norv
DT	16-OCT-2001 (Rel. 40, Last sequence update)	24.3	1237	1 B3A2_MOUSE	1	1237	1	P13808	mus musculus
DT	16-OCT-2001 (Rel. 40, Last annotation update)	24.3	1237	1 B3A2_RABIT	1	1237	1	P48746	oryctolagus
DE	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)	24.3	1241	1 B3A2_HUMAN	1	1241	1	P04920	homo sapien
DE	(AMPK gamma3).	24.3	2373	1 CCAH_HUMAN	1	2373	1	O95180	homo sapien
GN	PRKAG3 OR AMPKG3.	24.1	180	1 RSS_SYNTP6	1	180	1	O24705	synechococc
OS	Homo sapiens (Human).	24.1	200	1 NUJM_NEUCR	1	200	1	P25710	neurospora
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	24.1	363	1 CSP_PLAKH	1	363	1	P02894	plasmodium
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	24.1	478	1 GSR2_HUMAN	1	478	1	Q9ozm5	homo sapien
OX	NCBI_TaxID=9606;	24.1	1359	1 ATRX_CABEL	1	1359	1	Q9u7e0	caenorhabdi
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=20164049; PubMed=10698692;								
RA	Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;								
RT	"Characterization of AMP-activated protein kinase gamma-subunit								
RT	isoforms and their role in AMP binding.";								
RL	Biochem. J. 346:659-669(2000).								
RN	[2]								
RC	SEQUENCE FROM N.A.								
RX	TISSUE=Skeletal muscle;								
RA	MEDLINE=20280150; PubMed=10818001;								
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,								
RA	Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,								
RA	Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,								
RA	Andersson L.;								
RT	"A mutation in PRKAG3 associated with excess glycogen content in pig								
RT	skeletal muscle.";								
RL	Science 288:1248-1251(2000).								
CC	-!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID								
CC	SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO								
CC	REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND								
CC	INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-								
CC	SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE								
CC	IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.								
CC	-!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A								
CC	GAMMA NON-CATALYTIC REGULATORY SUBUNITS.								
CC	-!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART								
CC	AND PANCREAS.								
CC	-!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA								
CC	SUBUNIT FAMILY.								
CC	-!- SIMILARITY: CONTAINS 4 CBS DOMAINS.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL; AJ249977; CAB65117.1; ALT_INIT.								
DR	EMBL; AF214519; AAF73987.1; -.								

```
DR Genew: HGNC:9387; PRKAG3.
DR MW: 604976; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 147 201 CBS 1.
FT DOMAIN 228 282 CBS 2.
FT DOMAIN 303 356 CBS 3.
FT DOMAIN 375 428 CBS 4.
FT CONFLICT 58 58 T -> A (IN REF. 1).
FT CONFLICT 163 164 MQ -> IE (IN REF. 1).
FT CONFLICT 398 398 Q -> K (IN REF. 1).
FT CONFLICT 461 464 ALGA -> PSGPEK (IN REF. 1).
SQ SEQUENCE 464 AA; 51514 MW; 53985C2C77003A63 CRC64;

Query Match 98.1%; Score 210; DB 1; Length 464;
Best Local Similarity 97.6%; Pred. No. 1;le-16;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GKRRKALRWTRQKSVGEPPGGGEGPRSRPAAESTGLEA 41
   |||||
DB 26 GKRRKALRWTRQKSVGEPPGGGEGPRSRPAAESTGLEA 66

RESULT 2
AAKL_PIG
ID AAKL_PIG STANDARD; PRT; 464 AA.
AC Q9MYP4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)
DE (AMPK gamma3).
GN PRKAG3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT RN(-) GLN-200.
RC TISSUE-Skeletal muscle;
RX MEDLINE=20280150; PubMed=10818001;
RA Milan D., Jeon J.-T., Looft C., Robic A., Thelander M.,
RA Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,
RA Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,
RA Andersson L.;
RT "A mutation in PRKAG3 associated with excess glycogen content in pig
RT skeletal muscle.";
RL Science 288:1248-1251(2000).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPIASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE
CC IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -!- TISSUE SPECIFICITY: MUSCLE.
CC -!- DISEASE: DEFECTS IN PRKAG3 (RN-) ARE ASSOCIATED WITH EXCESS
CC GLYCOGEN CONTENT (ABOUT 70%) IN SKELETAL MUSCLE. THIS MUTATION
CC ORIGINATED IN THE HAMPSHIRE BREED PIGS AND HAS BENEFICIAL EFFECTS
CC ON MEAT CONTENT BUT DETRIMENTAL EFFECTS ON PROCESSING YIELD. THUS,
CC THIS MUTATION IS OF CONSIDERABLE ECONOMIC SIGNIFICANCE IN THE PIG
CC BREEDING INDUSTRY.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF214521; AAF73989.1; -.
DR EMBL: AF214520; AAF73988.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation.
FT DOMAIN 172 226 CBS 1.
FT DOMAIN 253 307 CBS 2.
FT DOMAIN 328 381 CBS 3.
FT DOMAIN 400 453 CBS 4.
FT VARIANT 200 200 R -> Q (IN RN-).
SQ SEQUENCE 464 AA; 51308 MW; 17638CB12A2BA9DF CRC64;

Query Match 66.4%; Score 142; DB 1; Length 464;
Best Local Similarity 68.3%; Pred. No. 5.3e-09;
Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GKRRKALRWTRQKSVGEPPGGGEGPRSRPAAESTGLEA 41
   |||||
DB 26 GDQGNKASRWTRQEDVEEGPPGREGPQSRPVAESTGQEA 66

RESULT 3
HNFA_MOUSE
ID HNFA_MOUSE STANDARD; PRT; 628 AA.
AC P22361;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver specific
DE transcription factor 1F-B1) (LFB1).
GN TCFL1 OR HNF1A OR HNF-1A OR HNF-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088607; PubMed=2263635;
RA Kuo C.-J., Conley P.B., Hsieh C.L., Francke U., Crabtree G.R.;
RT "Molecular cloning, functional expression, and chromosomal
RT localization of mouse hepatocyte nuclear factor 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9838-9842(1990).
CC -!- FUNCTION: REQUIRED FOR THE EXPRESSION OF SEVERAL LIVER SPECIFIC
CC GENES. BINDS TO THE INVERTED PALINDROME 5'-GTTAATNATTAAC-3'.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Liver.
CC -!- SIMILARITY: BELONGS TO THE HNF1 HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57966; AAA37821.1; -.
DR PIR: A39262; A39262.
DR HSSP: P15257; ILFB.
DR TRANSFAC: T01211; -.
DR MGD: MGI:98504; Tcf1.
DR InterPro: IPR001356; Homeobox.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW Activator; Trans-acting factor.
```



FT HELIX 261 273  
 FT TURN 274 275  
 SQ SEQUENCE 628 AA: 67213 MW: 8D26093308C86A52 CRC64;  
  
 Query Match 28.7%; Score 61.5; DB 1; Length 628;  
 Best Local Similarity 39.5%; Pred. No. 8.9;  
 Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;  
  
 QY 2 KRAKALWTRQKSVGEPPGGEGPRSRPAESTGL 39  
 : : | | | | | | | | | | | | | | | | | |  
 DB 271 RRKEAFRHLAMDYNGPPPGPGP-ALPAHSSPGL 307  
  
 RESULT 5  
 HNFA\_HUMAN  
 ID HNFA\_HUMAN STANDARD: PRT: 631 AA.  
 AC P20823; Q99861;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (liver specific  
 DE transcription factor LF-B1) (LFB1) (Transcription factor-1) (TCF-1).  
 GN TCF1 OR HNFA1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=91184801; PubMed=1707031;  
 RA Bach I., Gaicheva-Gargova Z., Mattei M.-G., Simon-Chazottes D.,  
 RA Guenet J.-L., Cereghini S., Yaniv M.;  
 RT "Cloning of human hepatic nuclear factor 1 (HNF1) and chromosomal  
 RT localization of its gene in man and mouse.";  
 RL Genomics 8:155-164(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94038905; PubMed=7900999;  
 RA Bach I., Yaniv M.;  
 RT "More potent transcriptional activators or a transdominant inhibitor  
 RT of the HNF1 homeoprotein family are generated by alternative RNA  
 RT processing.";  
 RL EMBO J. 12:4229-4242(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT MODY3 LEU-447.  
 RX MEDLINE=97100943; PubMed=8945470;  
 RA Vaxillaire M., Southam L., Cox R.D., Lathrop G.M., Boriraj V.V.,  
 RA Chen X., Cox N.J., Oda Y., Yano H., le Beau M.M., Yamada S.,  
 RA Nishigori H., Takeda J., Fajans S.S., Hattersley A.T., Iwasaki N.,  
 RA Hansen T., Pedersen O., Polonsky K.S., Turner R.C., Velho G.,  
 RA Chevre J.-C., Froguel P., Bell G.I.;  
 RT "Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-  
 RT onset diabetes of the young (MODY3).";  
 RL Nature 384:455-458(1996).  
 RN [4]  
 RP VARIANTS MODY3 ARG-107; TRP-131; MET-260 AND HIS-272.  
 RX MEDLINE=97309269; PubMed=9166684;  
 RA Glucksmann M.A., Lehto M., Tayber O., Scotti S., Berkemeier L.,  
 RA Pulido J.C., Wu Y., Nir W.-J., Fang L., Markel P., Munnelly K.D.,  
 RA Goranson J., Ortho M., Young B.M., Whitacre J.L., McMenimen C.,  
 RA Rosanman M., Tuomi T., Warram J., Forsblom C.M., Carlsson M.,  
 RA Rosenzweig J., Kennedy G., Duyk G.M., Krolewski A.S., Groop L.C.,  
 RA Thomas J.D.;  
 RT "Novel mutations and a mutational hotspot in the MODY3 gene.";  
 RL Diabetes 46:1081-1086(1997).  
 RN [5]  
 RP VARIANTS MODY3 H-12; Q-131; Q-205 AND C-263, AND VARIANT NIDDM D-191.  
 RX MEDLINE=97431548; PubMed=9287053;  
 RA Iwasaki N., Oda N., Ogata M., Hara M., Hinokio Y., Oda Y.,  
 RA Yamagata K., Kanematsu S., Ohgawara H., Omori Y., Bell G.I.;

"Mutations in the hepatocyte nuclear factor-1alpha/MODY3 gene in  
 Japanese subjects with early- and late-onset NIDDM.";  
 Diabetes 46:1504-1508(1997).  
 [6]  
 RN VARIANT NIDDM MET-254, AND VARIANTS LEU-27 AND ASN-487.  
 RP MEDLINE=97431550; PubMed=9287055;  
 RX Yamada S., Nishigori H., Oda H., Takahashi K., Kitano N.,  
 RA Morikawa A., Takeuchi T., Takeda J.;  
 RA "Mutations in the hepatocyte nuclear factor-1alpha gene (MODY3) are  
 RT not a major cause of late-onset NIDDM in Japanese subjects.";  
 RL Diabetes 46:1512-1513(1997).  
 RN [7]  
 RP VARIANTS IDDM HIS-272 AND GLY-583.  
 RX MEDLINE=97458988; PubMed=9313763;  
 RA Yamada S., Nishigori H., Oda H., Utsugi T., Yanagawa T., Maruyama T.,  
 RA Onigata K., Nagashima K., Nagai R., Morikawa A., Takeuchi T.,  
 RA Takeda J.;  
 RT "Identification of mutations in the hepatocyte nuclear factor (HNF)-1-  
 RT alpha gene in Japanese subjects with IDDM.";  
 RL Diabetes 46:1643-1647(1997).  
 RN [8]  
 RP VARIANTS MODY3, AND VARIANT ATYPICAL DIABETES SER-574.  
 RX MEDLINE=98052398; PubMed=9392505;  
 RA Boutin P., Chevre J.-C., Hani E.H., Gomis R., Pardini V.C.,  
 RA Guillausseau P.-J., Vaxillaire M., Velho G., Froguel P.;  
 RT "An automated fluorescent single-strand conformation polymorphism  
 RT technique for screening mutations in the hepatocyte nuclear  
 RT factor-1alpha gene (maturity-onset diabetes of the young).";  
 RL Diabetes 46:2108-2109(1997).  
 RN [9]  
 RP VARIANTS MODY3 GLN-131; GLN-229; GLY-241 AND HIS-272.  
 RX MEDLINE=97184312; PubMed=9032114;  
 RA Kaisaki P.J., Menzel S., Lindner T., Oda N., Rjasanowski I., Sahn J.,  
 RA Meincke G., Schulze J., Schmechel H., Petzold C., Ledermann H.M.,  
 RA Sachse G., Boriraj V.V., Menzel R., Kerner W., Turner R.C.,  
 RA Yamagata K., Bell G.I.;  
 RT "Mutations in the hepatocyte nuclear factor-1alpha gene in MODY and  
 RT early-onset NIDDM: evidence for a mutational hotspot in exon 4.";  
 RL Diabetes 46:528-535(1997).  
 RN [10]  
 RP VARIANTS MODY3 THR-129; TRP-131; TRP-159; LEU-519 AND ILE-620.  
 RX MEDLINE=97230240; PubMed=9075818;  
 RA Freyling T.M., Bulman M.P., Ellard S., Appleton M., Dronsfield M.J.,  
 RA Mackie A.D., Baird J.D., Kaisaki P.J., Yamagata K., Bell G.I.,  
 RA Bain S.C., Hattersley A.T.;  
 RT "Mutations in the hepatocyte nuclear factor-1alpha gene are a common  
 RT cause of maturity-onset diabetes of the young in the U.K.";  
 RL Diabetes 46:720-725(1997).  
 RN [11]  
 RP VARIANTS MODY3 ASN-128; TYR-143 AND LEU-447.  
 RX MEDLINE=97230241; PubMed=9075819;  
 RA Hansen T., Elberg H., Rouard M., Vaxillaire M., Holst J.J., Almnd K.,  
 RA Rasmussen S.K., Fridberg M., Urhammer S.A., Holst J.J., Almnd K.,  
 RA Ewald S.M., Hansen L., Bell G.I., Pedersen O.;  
 RT "Novel MODY3 mutations in the hepatocyte nuclear factor-1alpha gene:  
 RT evidence for a hyperexcitability of pancreatic beta-cells to  
 RT intravenous secretagogues in a glucose-tolerant carrier of a p447L  
 RT mutation.";  
 RL Diabetes 46:726-730(1997).  
 RN [12]  
 RP VARIANTS LEU-27; VAL-98 AND ASN-487.  
 RX MEDLINE=97278987; PubMed=9133564;  
 RA Urhammer S.A., Fridberg M., Hansen T., Rasmussen S.K., Moeller A.M.,  
 RA Clausen J.O., Pedersen O.;  
 RT "A prevalent amino acid polymorphism at codon 98 in the hepatocyte  
 RT nuclear factor-1alpha gene is associated with reduced serum C-peptide  
 RT and insulin responses to an oral glucose challenge.";  
 RL Diabetes 46:912-916(1997).  
 RN [13]  
 RP VARIANT NIDDM GLN-583, AND VARIANTS LEU-27; VAL-98 AND ASN-487.  
 RX MEDLINE=97266232; PubMed=9112026;  
 RA Urhammer S.A., Rasmussen S.K., Kaisaki P.J., Oda N., Yamagata K.,  
 RA Moeller A.M., Fridberg M., Hansen L., Hansen T., Bell G.I.;

RA Pedersen O.;  
 RT "Genetic variation in the hepatocyte nuclear factor-1 alpha gene in  
 RT Danish Caucasians with late-onset NIDDM.";  
 RL Diabetologia 40:473-475(1997).  
 RN [14]  
 RP VARIANTS MODY3 CVS-122; PHE-142 AND GLN-159.  
 RX MEDLINE-97252393; PubMed-9097962;  
 RA Vaxillaire M., Roudard M., Yamagata K., Oda N., Kaisaki P.J.,  
 RA Boriraj V.V., Chevre J.-C., Boccio V., Cox R.D., Lathrop G.M.,  
 RA Dussoix P., Philippe J., Timsit J., Charpentier G., Velho G.,  
 RA Bell G.I., Froquel P.;  
 RT "Identification of nine novel mutations in the hepatocyte nuclear  
 RT factor 1 alpha gene associated with maturity-onset diabetes of the  
 RT young (MODY3).";  
 RL Hum. Mol. Genet. 6:583-586(1997).  
 RN [15]  
 RP VARIANTS LEU-27; ASN-487 AND ARG-514.  
 RX MEDLINE-98265912; PubMed-9604876;  
 RA Behn P.S., Masson J., Chayen S., Smolovitch I., Thomas J.D.,  
 RA Glaser B., Permut M.A.;  
 RT "Hepatocyte nuclear factor 1alpha coding mutations are an uncommon  
 RT contributor to early-onset type 2 diabetes in Ashkenazi Jews.";  
 RL Diabetes 47:967-969(1998).  
 RN [16]  
 RP VARIANTS MODY3 ASP-31; TRP-159; THR-161; TRP-200 AND TRP-271.  
 RX MEDLINE-98425814; PubMed-9754819;  
 RA Chevre J.-C., Hani E.H., Boutin P., Vaxillaire M., Blanche H.,  
 RA Vionnet N., Pardini V.C., Timsit J., Largier E., Charpentier G.,  
 RA Beckers D., Maes M., Bellanne-Chantelot C., Velho G., Froquel P.;  
 RT "Mutation screening in 18 Caucasian families suggest the existence of  
 RT other MODY genes.";  
 RL Diabetologia 41:1017-1023(1998).  
 RN [17]  
 RP VARIANTS IDDM LYS-48 AND GLY-241.  
 RX MEDLINE-99083111; PubMed-9867222;  
 RA Moeller A.M., Dalggaard L.T., Pociot F., Nerup J., Hansen T.,  
 RA Pedersen O.;  
 RT "Mutations in the hepatocyte nuclear factor-1alpha gene in Caucasian  
 RT families originally classified as having type 1 diabetes.";  
 RL Diabetologia 41:1528-1531(1998).  
 RN [18]  
 RP VARIANTS MODY3 ARG-537 AND LYS-619.  
 RX MEDLINE-98289379; PubMed-9626139;  
 RA Elbein S.C., Teng K., Yount P., Scroggin E.;  
 RT "Linkage and molecular scanning analyses of MODY3/hepatocyte nuclear  
 RT factor-1 alpha gene in typical familial type 2 diabetes: evidence for  
 RT novel mutations in exons 8 and 10.";  
 RL J. Clin. Endocrinol. Metab. 83:2059-2065(1998).  
 RN [19]  
 RP VARIANTS LEU-27 AND ASN-487.  
 RX MEDLINE-98284534; PubMed-9621514;  
 RA Nishigori H., Yamada S., Kohama T., Utsugi T., Shimizu H.,  
 RA Takeuchi T., Takeda J.;  
 RT "Mutations in the hepatocyte nuclear factor-1 alpha gene 'MODY3' are  
 RT not a major cause of early-onset non-insulin-dependent 'type 2'  
 RT diabetes mellitus in Japanese.";  
 RL J. Hum. Genet. 43:107-110(1998).  
 RN [20]  
 RP VARIANTS MODY3 HIS-12; ASN-158; GLN-159 AND CVS-203.  
 RX MEDLINE-99176457; PubMed-10078571;  
 RA Yamada S., Tomura H., Nishigori H., Shio K., Mabe H., Iwatani N.,  
 RA Takumi T., Kito Y., Moriya N., Muroya K., Ogata T., Onigata K.,  
 RA Morikawa A., Inoue I., Takeda J.;  
 RT "Identification of mutations in the hepatocyte nuclear factor-1alpha  
 RT gene in Japanese subjects with early-onset NIDDM and functional  
 RT analysis of the mutant proteins.";  
 RL Diabetes 48:645-648(1999).  
 RN [21]  
 RP VARIANTS MODY3 GLU-117 AND TYR-143.

Query Match 28.7%; Score 61.5; DB 1; Length 631;  
 Best Local Similarity 39.5%; Pred. NO. 8.9;  
 Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTKQKSVGEPPGQGGPRSRPAESTGL 39  
 :|: :|: | | | | | :|: | | | | |  
 DB 271 RRKEAFRKLAMDYSGPPPGPGP-ALPAHSSPGL 307  
 RESULT 6  
 HIG\_DROME  
 ID HIG\_DROME STANDARD; PRT; 958 AA.  
 AC Q09101;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Locomotion-related protein Hikaru genki precursor.  
 GN HIG.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head;  
 RX MEDLINE-93213498; PubMed-8461133;  
 RA Hoshino M., Matsuzaki F., Nabeshima Y.-I., Hama C.;  
 RT "Hikaru genki, a CNS-specific gene identified by abnormal locomotion  
 RL in Drosophila, encodes a novel type of protein.";  
 RL Neuron 10:395-407(1993).  
 CC -!- FUNCTION: HAS A ROLE IN THE DEVELOPMENT OF CNS FUNCTIONS INVOLVED  
 CC IN LOCOMOTOR ACTIVITY.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1, 2, 3 (SHOWN HERE) AND 4; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PCC NEURONS AND NEUROBLASTS  
 CC IN THE PROCEPHALIC NEUROGENIC REGION IN THE CENTRAL NERVOUS  
 CC SYSTEM.  
 CC -!- DEVELOPMENTAL STAGE: MOST ABUNDANT DURING AND/OR AFTER NEURONAL  
 CC DIFFERENTIATION AND DURING CELL SPECIFICATION OR AXOGENESIS.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUGLOBULIN SUPERFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 IMMUGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D13884; BAA02984.1; -;  
 DR EMBL; D13885; BAA02985.1; -;  
 DR EMBL; D13886; BAA02986.1; -;  
 DR EMBL; D13887; BAA02987.1; -;  
 DR HSP; P10998; LVVD.  
 DR FlyBase; FBgn0010114; hig.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00084; sushi; 5.  
 DR SMART; SM00032; CCP; 5.  
 DR SMART; SM00409; IG; 1.  
 KW Glycoprotein; Alternative splicing; Immunoglobulin domain; Repeat;  
 KW SUSHI; Signal.  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 958 LOCOMOTION-RELATED PROTEIN HIKARU GENKI.  
 FT DOMAIN 630 709 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 713 769 SUSHI 1.  
 FT DOMAIN 772 828 SUSHI 2.  
 FT DOMAIN 831 891 SUSHI 3.  
 FT DOMAIN 893 952 SUSHI 4.  
 FT SITE 318 320 CELL ATTACHMENT SITE.

Pfam: PF00036; ehand; 3.  
DR PRINTS; PR01362; CALFLAGIN.  
DR ProDom; PD000012; EF-hand; 2.  
DR SMART; SM00054; EFh; 3.  
DR PROSITE; PS00018; EF\_HAND; 3.  
KW Calcium-binding; Repeat.  
FT CA\_BIND 142 153 EF-HAND 1 (POTENTIAL).  
FT DOMAIN 178 189 ANCESTRAL CALCIUM SITE 2 (POTENTIAL).  
FT CA\_BIND 219 230 EF-HAND 3 (POTENTIAL).  
FT CA\_BIND 256 267 EF-HAND 4 (POTENTIAL).  
SQ SEQUENCE 275 AA: 30433 MW; 1AEB49C93AD67DB7 CRC64;

Query Match 27.6%; Score 59; DB 1; Length 275;  
Best Local Similarity 29.5%; Pred.No. 7.5;  
Matches 13; Conservative 7; Mismatches 12; Indels 12; Gaps 1;

QY 3 RRAKALEWTRQKSVEEGE-----PPGQGEPGRSRPA 34  
:: : || :||| :: || || |||| | |  
DB 43 KKRGUGRSKRKTGSSEQTGPEAPGSSNNPFSTGEGPAGAP 86

RESULT 8  
SYNL\_HUMAN SYNL\_HUMAN STANDARD; PRT; 705 AA.

ID ID SYNL\_HUMAN STANDARD; PRT; 705 AA.  
AC PI7600; O75825;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Synapsin I (Brain protein 4.1).  
GN SYNL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90243651; PubMed=2110562;  
RA Suedhof T.C.;  
RT "The structure of the human synapsin I gene and protein.";  
RL J. Biol. Chem. 265:7849-7852(1990).  
RN [2]  
RP SEQUENCE OF 1-125 FROM N.A.  
RX MEDLINE=90368667; PubMed=2118519;  
RA Sauerwald A., Hoessche C., Oschwald R., Killmann M.W.;  
RT "The 5'-flanking region of the synapsin I gene. A G-C-rich, TATA- and CAT-less, phylogenetically conserved sequence with cell type-specific promoter function.";  
RL J. Biol. Chem. 265:14932-14937(1990).  
RN [3]  
RP SEQUENCE OF 1-258 FROM N.A.  
RA Grafham D.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
CC -! FUNCTION: NEURONAL PHOSPHOPROTEIN THAT COATS SYNAPTIC VESICLES,  
CC BINDS TO THE CYTOSKELETON, AND IS BELIEVED TO FUNCTION IN THE  
CC REGULATION OF NEUROTRANSMITTER RELEASE.  
CC -! SUBCELLULAR LOCATION: SYNAPSE.  
CC -! ALTERNATIVE PRODUCTS: 2 ISOFORMS; SYNAPSIN IA (SHOWN HERE) AND  
CC SYNAPSIN IB; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -! PTM: SUBSTRATE OF AT LEAST FOUR DIFFERENT PROTEIN KINASES. IT IS  
CC PROBABLE THAT PHOSPHORYLATION PLAYS A ROLE IN THE REGULATION OF  
CC SYNAPSIN I IN THE NERVE TERMINAL.  
CC -! SIMILARITY: BELONGS TO THE SYNAPSIN FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: M58378; AAC41910.1; -  
CC DR



```
CC EMBL; X90857; CAA62368.1; -.
DR EMBL; Z69666; CAA93521.1; -.
DR MIM; 600928; -.
DR InterPro: IPR0005365; UPF0171.
DR Pfam; PF03666; UPF0171.1.
SQ SEQUENCE 569 AA; 63604 MW; 44BEF42AA7F2841D CRC64;

Query Match      25.9%; Score 55.5; DB 1; Length 569;
Best Local Similarity 46.2%; Pred. No. 38;
Matches 18; Conservative 3; Mismatches 15; Indels 3; Gaps 2;

Qy 1 KRRRAKAL-RWTQKSVESGEPGCGEGRSRPRAESTG 38
Db 16 GSRGNKLLFRYPQRSQE--HPASQTSKPSRYAASNTG 52

RESULT 11
Z145_HUMAN
ID Z145_HUMAN STANDARD; PRT; 673 AA.
AC Q05516;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein PLZF (Promyelocytic leukemia zinc finger
DE protein) (zinc finger protein 145).
GN ZNF145 OR PLZF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart ventricle;
RA Chen S.-J., Zellent A., Tong J.-H., Yu H.-Q., Wang Z.-Y., Derre J.,
RA Berger R., Waxman S., Chen Z.;
RA "Rearrangements of the retinoic acid receptor alpha and promyelocytic
RA leukemia zinc finger genes resulting from t(11;17)(q23;q21) in a
RT patient with acute promyelocytic leukemia.";
RL J. Clin. Invest. 91:2260-2267(1993).
RN [2]
RP SEQUENCE OF 424-455 FROM N.A.
RX MEDLINE=93253074; PubMed=8387545;
RA Chen S.-J., Zellent A., Tong J.-H., Yu H.-Q., Wang Z.-Y., Derre J.,
RA Berger R., Waxman S., Chen Z.;
RT "Rearrangements of the retinoic acid receptor alpha and promyelocytic
RT leukemia zinc finger genes resulting from t(11;17)(q23;q21) in a
RT patient with acute promyelocytic leukemia.";
RL J. Clin. Invest. 91:2260-2267(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 7-122.
RX MEDLINE=20005701; PubMed=10537309;
RA Li X., Peng H., Schultz D.C., Lopez-Guisa J.M., Rauscher F.J. III,
RA Marmorstein R.;
RT "Structure-function studies of the WT/POZ transcriptional repression
RT domain from the promyelocytic leukemia zinc finger oncoprotein.";
RL Cancer Res. 59:5275-5282(1999).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR. MAY PLAY A ROLE IN
CC MYELOID MATURATION AND IN THE DEVELOPMENT AND/OR MAINTENANCE OF
CC OTHER DIFFERENTIATED TISSUES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: PLFZA AND PLZFB (SHOWN HERE);
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: WITHIN THE HEMATOPOIETIC SYSTEM, PLZF IS
CC EXPRESSED IN BONE MARROW, EARLY MYELOID CELL LINES AND PERIPHERAL
CC BLOOD MONONUCLEAR CELLS. ALSO EXPRESSED IN THE OVARY, AND AT LOWER
CC LEVELS, IN THE KIDNEY AND LUNG.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- DISEASE: A FORM OF ACUTE PROMYELOCYTIC LEUKEMIA (APL) IS
CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(11;17)(q32;q21)
CC WHICH INVOLVES ZNF145 AND RETINOIC ACID RECEPTOR ALPHA (RARA).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
```

```
CC FINGER PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z19002; CAA79489.1; -.
DR EMBL; S60093; AAC60590.2; -.
DR PDB; 1CS3; 09-AUG-00.
DR TRANSFAC; T02336; -.
DR Genew; HGNC:12930; ZNF145.
DR MIM; 176797; -.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 9.
DR Pfam; PF00651; BTB; 1.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00355; Znf_C2H2; 9.
DR PROSITE; PS00097; BTB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 9.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat; Chromosomal translocation; Proto-oncogene;
KW Phosphorylation; Alternative splicing; 3D-structure.
FT DOMAIN 34 96
FT ZINC_FINGERS.
FT ZN_FING 404 652
FT ZN_FING 404 426
FT ZN_FING 432 454
FT ZN_FING 461 483
FT ZN_FING 490 512
FT ZN_FING 518 540
FT ZN_FING 546 568
FT ZN_FING 574 596
FT ZN_FING 602 624
FT ZN_FING 630 652
FT SITE 394 395
FT MOD_RES 76 76
FT MOD_RES 184 184
FT MOD_RES 197 197
FT MOD_RES 256 256
FT MOD_RES 282 282
FT MOD_RES 528 528
FT MOD_RES 555 555
FT VARSPLIC 255 377
SQ SEQUENCE 673 AA; 74332 MW; 7CD7319E2A32109D CRC64;

Query Match      25.9%; Score 55.5; DB 1; Length 673;
Best Local Similarity 37.9%; Pred. No. 45;
Matches 11; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

Qy 5 AKALFWTRQKSVESGEPGQ-GEGRSRP 32
Db 292 ARELHYGRESAQVPPAEAGQAPTGRP 320

RESULT 12
SLBP_HUMAN
ID SLBP_HUMAN STANDARD; PRT; 270 AA.
AC Q14493;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Histone RNA hairpin-binding protein (Histone stem-loop binding
DE protein).
GN HBP OR SLBP.
OS Homo sapiens (Human).
```





DR	Pfam; PF02078; Synapsin; 1.
DR	Pfam; PF02750; Synapsin_C; 1.
DR	PRINTS; PROI368; SYNAPSIN.
DR	PROSITE; PS00415; SYNAPSIN_1; 1.
DR	PROSITE; PS00416; SYNAPSIN_2; 1.
KW	Synapse; Phosphorylation; Nerve; Neuron; Repeat; Actin-binding;
KW	Alternative splicing.
FT	DOMAIN 1 28 A.
FT	DOMAIN 29 112 B (LINKER).
FT	DOMAIN 113 420 C (ACTIN-BINDING AND SYNAPTIC-VESICLE BINDING).
FT	DOMAIN 421 555 D (PRO-RICH LINKER).
FT	DOMAIN 556 704 E.
FT	MOD_RES 9 9 PHOSPHORYLATION (BY PKA AND CAMK1).
FT	MOD_RES 566 566 PHOSPHORYLATION (BY CAMK2).
FT	MOD_RES 603 603 PHOSPHORYLATION (BY CAMK2).
FT	VARSPLIC 660 668 NKQSLNNA -> KASPAQAQP (IN SYNAPSIN IB).
FT	VARSPLIC 669 704 MISSING (IN SYNAPSIN IB).
SQ	SEQUENCE 704 AA; 73988 MW; 65799FEF7CFE1B85 CRC64;
Query Match	25.7%; Score 55; DB 1; Length 704;
Best Local Similarity	42.3%; Pred. No. 54;
Matches 11; Conservative	4; Mismatches 11; Indels 0; Gaps
QY	13 QKSVEEGEPFGQGGRSPRAAESTG 38   :   :   :   :
Dd	514 QQSAQTMTQQGQRSPVAGPG 539
RESULT 15	
ID	GUNA_THIEFU STANDARD; PRT: 880 AA.
AC	P26221; O08167;
DT	01-MAY-1992 (Rel. 22, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Endoglucanase E-4 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase E-4) (Cellulase E-4) (Cellulase B4).
DE	CELD.
OS	Thermomonospora fusca.
OC	Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC	Actinomycetales; Streptosporangineae; Nocardiopsaceae; Thermobifida;
OX	NCB TaxID=2021;
RX	[1]
RC	SEQUENCE FROM N.A.
RC	STRAIN-YX:
RX	MEDLINE=94028932; PubMed=8215374;
Jung E.D., Lao G., Irwin D., Barr B.K., Benjamin A., Wilson D.B.;	
"DNA sequences and expression in Streptomyces lividans of an exoglucanase gene and an endoglucanase gene from Thermomonospora fusca."	
Appl. Environ. Microbiol. 59:3032-3043(1993).	
[2]	
RP REVISIONS.	
RA Wilson D.B.;	
RA Submitted (FEB-1997) to the EMBL/GenBank/DDJ databases.	
[3]	
RP PARTIAL SEQUENCE FROM N.A.	
RC STRAIN-YX:	
RX MEDLINE=91258320; PubMed=1904434;	
Lao G., Chngas G.S., Jung E.D., Wilson D.B.;	
"DNA sequences of three beta-1,4-endoglucanase genes from Thermomonospora fusca;"	
J. Bacteriol. 173:3397-3407(1991).	
[4]	
RP SEQUENCE OF 47-67.	
RA Wilson D.B.;	
RA "Cellulases of Thermomonospora fusca.";	
RL Meth. Enzymol. 160:314-323(1988).	
[5]	
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 47-651.	
RX MEDLINE=97475222; Pubmed=9334746;	
Sakon J., Irwin D., Wilson D.B., Karplus P.A.;	

```
KT "Structure and mechanism of endo/exocellulase E4 from Thermomonospora
RL fusca.";
CC Nat. Struct. Biol. 4:810-818(1997).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose.
CC -!- PATHWAY: Cellulose degradation.
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL
CC HYDROLASES).
CC -!- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20093; AAB42155.1; -.
DR EMBL; M73322; AAA27397.1; ALT_SEQ.
DR PIR; B42360; B42360.
DR PDB; LJS4; 17-SEP-97.
DR PDB; 1TF4; 04-SEP-97.
DR PDB; 3TF4; 04-SEP-97.
DR PDB; 4TF4; 04-SEP-97.
DR InterPro; IPR001919; Bac_cellose-bind.
DR InterPro; IPR001956; CBD_3.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FN_III_repeat.
DR InterPro; IPR001701; GH_9.
DR InterPro; IPR001230; Prenyl_site.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00553; CBM_2; 1.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR Pfam; PF00942; CBM_3; 1.
DR PRINTS; PR00014; FN1YPEIII.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS00561; CBD_BACTERIAL; 1.
DR PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
DR PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 46
FT CHAIN 47 880 ENDOGLUCANASE E-4.
FT DOMAIN 776 880 CELLULOSE-BINDING.
FT ACT_SITE 427 427 BY SIMILARITY.
FT ACT_SITE 461 461 BY SIMILARITY.
FT ACT_SITE 470 470 BY SIMILARITY.
SQ SEQUENCE 880 AA; 95202 MW; 5EA9A6ABF45A4D9A CRC64;

Query Match 25.7%; Score 55; DB 1; Length 880;
Best Local Similarity 83.3%; Pred No. 67;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 17 EGEPPPGGEGP 28
Db 652 EGEPPGGGEGP 663

Search completed: June 6, 2003, 11:01:30
Job time : 3.50788 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:58:04 ; Search time 5.31349 Seconds  
(without alignments)  
1589.904 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GKRAKALRWTRQKSEVERGE.....PGQGGRSRPAAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_fodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	30.4	399	11 Q8VCQ7	Q8vcq7 mus musculus
2	62	29.0	493	10 Q81782	Q81782 arabidopsis
3	61.5	28.7	866	5 Q917E3	Q917e3 drosophila
4	61.5	28.7	933	5 Q917E4	Q917e4 drosophila
5	61.5	28.7	958	5 Q9V560	Q9v560 drosophila
6	60	28.0	136	10 Q42174	Q42174 arabidopsis
7	60	28.0	314	10 Q8S835	Q8s835 oryza sativ
8	60	28.0	486	10 Q9LQJ7	Q9luj7 arabidopsis
9	60	28.0	793	4 Q9H522	Q9h522 homo sapien
10	59.5	27.8	330	10 Q8S5D8	Q8s5d8 oryza sativ
11	59.5	27.8	1400	5 Q9NKS7	Q9nks7 leishmania
12	59	27.6	514	5 Q961A3	Q961a3 drosophila
13	59	27.6	863	5 Q9VRV3	Q9vr3 drosophila
14	58.5	27.3	207	15 Q8US87	Q8us87 human immun
15	58	27.1	277	16 Q9LQD4	Q9lqd4 streptomyce
16	58	27.1	303	16 Q98GW6	Q98gw6 rhizobium l

Q8qy66 mus musculu  
Q94lu8 oryza sativ  
Q9as24 oryza sativ  
Q69020 human herpe  
Q8sq03 canis famil  
Q8yb61 brucella me  
Q95da7 caenorhabdi  
Q95qa6 caenorhabdi  
Q95qa8 caenorhabdi  
Q9rkc5 streptomyce  
Q9sn19 oryza sativ  
Q9n7j7 homo sapien  
Q12845 homo sapien  
Q8x0s8 neurospora  
Q9lch6 macropodid  
Q8vlj6 ralstonia s  
Q8qx57 ralstonia s  
Q90we6 leucopsario  
Q9ab03 caulobacter  
Q8vxg7 zea mays (m  
Q9rzi5 deinococcus  
Q9aun1 oryza sativ  
Q9awx0 oryza sativ  
Q8s207 oryza sativ  
Q9cqs4 mus musculu  
Q8ty27 methanopyru  
Q93jh7 streptomyce  
Q8tal4 homo sapien  
Q925a9 streptomyce

399 27.1 11 Q9QY66  
516 10 Q94IU8  
88 10 Q9AS24  
198 12 Q69020  
299 6 Q8SQ03  
565 16 Q8YB61  
603 5 Q95DA7  
663 5 Q95QA6  
706 5 Q95QA8  
420 16 Q9RKC5  
952 10 Q9SN19  
278 4 Q9H7J7  
278 4 Q12845  
367 3 Q8X0S8  
1234 12 Q91CH6  
246 16 Q8VLJ6  
304 16 Q8QX57  
326 13 Q90WE6  
435 16 Q9ABD3  
703 10 Q8VXG7  
969 16 Q9RZI5  
254 10 Q9AUN1  
266 10 Q9AWX0  
287 10 Q8S207  
418 11 Q9CQS4  
447 17 Q8TY27  
618 16 Q93JH7  
673 4 Q8TAL4  
943 16 Q925A9

#### ALIGNMENTS

#### RESULT 1

Q8VCQ7 PRELIMINARY; PRT; 399 AA.

AC Q8VCQ7;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Open reading frame 6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC019440; AAH19440.1; -  
SQ SEQUENCE 399 AA; 42995 MW; B5CA88342B56932B CRC64;

Query Match 30.4%; Score 65; DB 11; Length 399;

Best Local Similarity 47.1%; Pred. No. 3.2;

Matches 16; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 3 RRAKALRWTRQKSEVERGEPPQGGGRSRPAAES 36

DB 63 RLALRLRLREQREAEDPEEAGLGRPRPAGGLS 96

#### RESULT 2

Q81782 PRELIMINARY; PRT; 493 AA.

ID Q81782;  
AC Q81782;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-OCR-2000 (TREMBLrel. 15, Last annotation update)

DE Hypothetical 56.2 kDa protein.

GN F28M20.160 OR AT4G31650.

OS Arabidopsis thaliana (Mouse-ear cress).





RC	STRAIN-COLUMBIA;
RL	Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RM	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.A.
RC	STRAIN-COLUMBIA;
RX	MEDLINE=20277480; PubMed=10819329;
RA	Nakamura Y.;
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT	features of the regions of 4,504,864 bp covered by sixty Pl and TAC
RT	clones.";
RNA	DNA Res. 7:131-135(2000).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA	Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA	Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA	Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA	Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA	Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA	Ecker J.R.;
RT	"Arabidopsis cDNA clones";
RN	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN	[4]
RP	SEQUENCE FROM N.A.
RA	Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
RA	Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA	Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA	Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA	Ecker J.R.;
RT	"Arabidopsis ORF clones";
RT	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DL	EMBL; AB021223; BAB01239.1; -
DR	EMBL; AY058085; AAL24193.1; -
DR	EMBL; AY090307; AAL90968.1; -
DR	HSSP; P02853; 2PHL
DR	InterPro; IPR001064; Crystallin.
DR	InterPro; IPR001113; Seedstore_7s.
DR	Pfam; PF00546; Seedstore_7s; 1.
DR	Pfam; PF02808; Seedstore_7s; C. 1.
DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN 1
SQ	SEQUENCE 486 AA; 55063 MW; 325ECF68D9A6345B CRC64;
Query Match            56.5%;    Score 60;    DB 10;    Length 486;	
Best Local Similarity    56.5%;    Pred. No. 17;	
Matches 13;    Conservative         2;    Mismatches         6;    Indels         2;    Gaps         1;	
QY	16 VEEGEPGGQ--EGPRSPPAES 36 ::
Db	29 IDEFEPPQGEQEGRRRPPGGGS 51 ::
RESULT 9	
Q9H5Z2	
ID	Q9H5Z2 PRELIMINARY; PRT; 793 AA.
DC	Q9H5Z2;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	CDNA: FLJ22757 fis, clone KAI0803.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=ILEAL MUCOSA;
RA	Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA	Nakajima Y., Nakano T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,

```
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isodai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK026410; BAB15475.1; -.
DR InterPro: IPR005112; dENN.
DR InterPro: IPR001194; dENN.
DR InterPro: IPR005113; uDENN.
DR Pfam: PF03455; dENN; 1.
DR Pfam: PF02141; dENN; 1.
DR Pfam: PF03456; uDENN; 1.
DR SEQUENCE 793 AA; 86230 MW; 3F37ACA836F24AF4 CRC64;

Query Match 28.0%; Score 60; DB 4; Length 793;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 16; Conservative 4; Mismatches 8; Indels 4; Gaps 2;

QY 1 KRRAKALRWTRQKSVREG--PPQGEGGPRS 30
Db 504 GKNR--PLRPSRRRLQEGTSPPGAGTPPLS 533

RESULT 10
Q8S5D8 PRELIMINARY; PRT; 330 AA.
AC Q8S5D8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 37.7 kDa protein.
GN OSJNB0047B19.26.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB0047B19, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC113339; AA08880.1; -.
KW Hypothetical protein.
SQ SEQUENCE 330 AA; 37674 MW; 01AD287515E7DD43 CRC64;

Query Match 27.8%; Score 59.5; DB 10; Length 330;
Best Local Similarity 37.2%; Pred. No. 13;
Matches 16; Conservative 5; Mismatches 13; Indels 9; Gaps 2;

QY 2 KRRAKALRWTRQKSVREGPPQGEGPRS---RPAESTGLEA 41
Db 38 RRRFAALWRR-----GGPPVTARGPQATARQPLATMRGSA 74

RESULT 11
Q9NKS7 PRELIMINARY; PRT; 1400 AA.
AC Q9NKS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE L509.6.
GN L509.6.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDELIN;
RA Myler P.J.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005836; AAF3594.1; -.
DR SEQUENCE 1400 AA; 149577 MW; 0CDCF4E925BEAEF4 CRC64;

Query Match 27.8%; Score 59.5; DB 5; Length 1400;
Best Local Similarity 36.4%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 11; Indels 17; Gaps 1;

QY 12 ROKSVEGEP-----PQQEGGPRSPAAESTG 38
Db 432 RQPEEEEEPRRAFLLKDDAVLAEGVQPADGEAAAGRAPAASSTG 475

RESULT 12
Q961A3 PRELIMINARY; PRT; 514 AA.
AC Q961A3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE LD26701P.
GN CG10274.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY051740; AAK93164.1; -.
DR FlyBase; FBgn0035690; CG10274.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000130; Zn_MTPeptdse.
DR Pfam: PF00096; zf-C2H2; 9.
DR PROSITE; PS00028; ZINC_FINGER_C2H2.1; UNKNOWN_7.
DR PROSITE; PS0157; ZINC_FINGER_C2H2.2; 9.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 514 AA; 59029 MW; 42DC53A6A4EC5F49 CRC64;

Query Match 27.6%; Score 59; DB 5; Length 514;
Best Local Similarity 35.9%; Pred. No. 24;
Matches 14; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 2 KRRAKALRWTRQKSVREGPPQGEGPRSPPAAESTGLE 40
Db 183 KRVLRAIETYRQVQEMGETPGEGFDNPPAPPVEGISIE 221

RESULT 13
Q9VRV3 PRELIMINARY; PRT; 863 AA.
AC Q9VRV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG10274 protein.
GN CG10274.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
```





```
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL161803; CAB82076.1; -.
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004044; KH_TYPE_2.
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00013; KH-domain; 1.
DR Pfam: PF00189; Ribosomal_S3_C; 1.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
DR SMART: SM00322; KH; 1.
DR TIGRFAMS: TIGR01009; rpsC_bact; 1.
DR PROSITE: PS00823; KH_TYPE_2; 1.
DR PROSITE: PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein.
SQ SEQUENCE 277 AA; 30273 MW; 5831536E64018135 CRC64;

Query Match 27.1%; Score 58; DB 16; Length 277;
Best Local Similarity 41.5%; Pred. No. 17;
Matches 17; Conservative 5; Mismatches 15; Indels 4; Gaps 2;

QY 1 CKRRKALRWTFQKSVVEGEPGQEGCPRSPRAAESTGLEA 41
Db |:|| | :|:| | | : | | : ||||| |
241 GERRG---RKPOQSAPAAEAP-KAEAPAAAAPAEETGTEA 277
```

Search completed: June 6, 2003, 11:02:51  
Job time : 7.31349 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 2.44133 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GKRAKALRWTKSVESGE.....PCQCGPRSRPAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61.5	28.7	415	4	US-08-927-219-8
2	61.5	28.7	630	4	US-08-927-219-2
3	61.5	28.7	630	4	US-08-927-219-4
4	61.5	28.7	631	4	US-08-927-219-127
5	60	28.0	187	4	US-09-199-637A-287
6	59	27.6	696	3	US-08-906-865-4
7	59	27.6	696	4	US-09-129-668-4
8	52	24.3	2353	4	US-08-984-709A-50
9	51	23.8	499	2	US-08-820-170A-40
10	51	23.8	499	3	US-09-055-699-40
11	51	23.8	499	4	US-09-273-565-40
12	51	23.8	499	4	US-09-565-538-40
13	51	23.8	499	4	US-09-661-468-40
14	51	23.8	503	4	US-09-562-737-62
15	51	23.8	1251	5	PCI-US95-02251-3
16	51	23.8	1252	1	US-08-199-780-3
17	51	23.8	1252	2	US-08-316-650-3
18	51	23.8	1253	3	US-08-479-722B-4
19	50.5	23.6	223	4	US-09-206-676C-1
20	50.5	23.6	226	4	US-09-206-676C-2
21	50.5	23.6	1497	1	US-08-623-679-7
22	50.5	23.6	1497	3	US-08-933-774-7
23	50.5	23.6	1497	4	US-09-181-030-7
24	50.5	23.6	1497	4	US-09-534-242-7
25	50.5	23.6	1497	4	US-09-454-854-7
26	50.5	23.6	1497	4	US-09-164-671-7
27	50.5	23.6	1533	1	US-08-623-679-9

28	50.5	23.6	1533	3	US-08-933-774-9	Sequence 9, Appli
29	50.5	23.6	1533	4	US-09-181-030-9	Sequence 9, Appli
30	50.5	23.6	1533	4	US-09-534-242-9	Sequence 9, Appli
31	50.5	23.6	1533	4	US-09-454-854-9	Sequence 9, Appli
32	50.5	23.6	1533	4	US-09-164-671-9	Sequence 9, Appli
33	50	23.4	410	1	US-07-945-283-4	Sequence 4, Appli
34	50	23.4	449	1	US-09-041-075A-21	Sequence 21, Appli
35	50	23.4	4928	4	US-09-036-987A-5	Sequence 5, Appli
36	50	23.4	4928	4	US-09-370-700-5	Sequence 5, Appli
37	49.5	23.1	106	4	US-08-785-065-5	Sequence 5, Appli
38	49.5	23.1	106	4	US-09-151-412-5	Sequence 5, Appli
39	49.5	23.1	233	2	US-08-458-568A-4	Sequence 4, Appli
40	49.5	23.1	448	4	US-09-310-463-18	Sequence 18, Appli
41	49.5	23.1	448	4	US-08-842-248A-18	Sequence 18, Appli
42	49	22.9	595	4	US-09-370-838-187	Sequence 187, App
43	49	22.9	882	4	US-09-413-814-78	Sequence 78, Appli
44	49	22.9	1780	1	US-08-769-309A-5	Sequence 5, Appli
45	49	22.9	1780	3	US-08-994-570-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-927-219-8  
; Sequence 8, Application US/08927219  
; Patent No. 6187533  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Graeme I.  
; APPLICANT: Yamagata, Kazuya  
; APPLICANT: Oda, Naohisha  
; APPLICANT: Katsaki, Pamela J.  
; APPLICANT: Furuta, Hiroto  
; APPLICANT: Horikawa, Yukio  
; APPLICANT: Menzel, Stephen  
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY  
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,219  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,679  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/028,056  
; FILING DATE: 02-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/025,719  
; FILING DATE: 10-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Willson, Mark B.  
; REGISTRATION NUMBER: 37,259  
; REFERENCE/DOCKET NUMBER: ARCD:272  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:

```

Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;
QY 2 KRRAKALRWTRQKSVEEGPPGGGEGPRSRPAESTGL 39
Db 271 KRKEEAFRKLAMDTYSGPPPGPGGP-ALPAHSSPGL 307

RESULT 3
US-08-927-219-4
; Sequence 4, Application US/08927219
; Patent No. 6187533
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme I.
; APPLICANT: Yamagata, Kazuya
; APPLICANT: Oda, Naohisha
; APPLICANT: Kaisaki, Pamela J.
; APPLICANT: Furuta, Hiroto
; APPLICANT: Horikawa, Yukio
; APPLICANT: Menzel, Stephen
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA
; TITLE OF INVENTION: AND HNF-4ALPHA
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,219
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,679
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,056
; FILING DATE: 02-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,719
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:272
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-927-219-4

Query Match 28.7%; Score 61.5; DB 4; Length 630;
Best Local Similarity 39.5%; Pred. No. 3.6;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;
QY 2 KRRAKALRWTRQKSVEEGPPGGGEGPRSRPAESTGL 39
Db 271 KRKEEAFRKLAMDTYSGPPPGPGGP-ALPAHSSPGL 307

RESULT 4

```

```
US-08-927-219-127
; Sequence 127, Application US/08927219
; Patent No. 6187533
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme I.
; APPLICANT: Yamagata, Kazuya
; APPLICANT: Oda, Naohisa
; APPLICANT: Kaisaki, Pamela J.
; APPLICANT: Furuta, Hiroto
; APPLICANT: Horikawa, Yukio
; APPLICANT: Menzel, Stephen
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA
; TITLE OF INVENTION: AND HNF-4ALPHA
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 60/927,219
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,679
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,056
; FILING DATE: 02-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,719
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: AHC0-272
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 127:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-927-219-127
Query Match 28.7%; Score 61.5; DB 4; Length 631;
Best Local Similarity 39.5%; Pred. No. 3.6;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRRAKALRWTRKSVVEGPPGCGEPRSRPAESTGL 39
   :|:| | | | | | | | | | | | | | | | |
Db 271 RRKEAPRHKLAWDYGPPPGPGP-ALPAHSSPGL 307

RESULT 5
US-09-199-637A-287
; Sequence 287, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
```

```
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRUS-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-287
Query Match 28.0%; Score 60; DB 4; Length 187;
Best Local Similarity 34.9%; Pred. No. 1.4;
Matches 15; Conservative 6; Mismatches 16; Indels 6; Gaps 1;

QY 1 GKRRAKALRWTRQ-----KSVVEGPPGCGEPRSRPAEST 37
   :|:| | | | | | | | | | | | | | | | |
Db 47 GRRANGARWTRRLPPRGRSLADAPAPCAASRARPRASST 89

RESULT 6
US-08-906-865-4
; Sequence 4, Application US/08906865
; Patent No. 6040168
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,865
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: /desc = "Synapsin Ia"
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
```





; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-661-468-40

Query Match 23.8%; Score 51; DB 4; Length 499;
Best Local Similarity 38.7%; Pred. No. 66;
Matches 12; Conservative 1; Mismatches 16; Indels 2; Gaps 1;

QY 10 WTRQKSVGEPPGGGPRSRPAESTGLE 40
| | | | | | | | | | | | | | | |
DB 21 WEHSKEVSEAPGGGGSGDGP--EESGQE 49

RESULT 14
US-09-562-737-62
; Sequence 62, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-62

Query Match 23.8%; Score 51; DB 4; Length 503;
Best Local Similarity 23.6%; Pred. No. 67;
Matches 17; Conservative 8; Mismatches 15; Indels 32; Gaps 2;

QY 2 KRRAKALRWTR-----KQSVSEGEPPGCGEGPK-- 29
|:::|::| |
DB 144 KKKSQAMRWRTVGQAFYCHKLSLQHTAQNADQGDSDRNSDGGSGEGPQLTGAEFV 203
| | | | | | | | | | | | | | | |
QY 30 SRPAESTGLEA 41
| | | | | | | | | | | | | | | |
DB 204 STATAETGIDA 215

RESULT 15
PCT-US95-02251-3
; Sequence 3, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UMIC009P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0824
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1251 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-02251-3

Query Match 23.8%; Score 51; DB 5; Length 1251;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 20 EPPGQGGGPRSRPAA 34
|:::|::|::|
DB 172 DPPGGGEGPPAQHAA 186

Search completed: June 6, 2003, 11:04:14
Job time : 3.44133 secs



GenCore version 5.1.6  
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 11:02:59 ; Search time 3.37478 Seconds  
(without alignments)  
1254.259 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91  
Perfect score: 214  
Sequence: 1 GKRRAKALRWTRQKSVEGE.....PQGGPRSRPAAESTGLEA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	214	100.0	489	10	US-09-826-581-6
2	60	28.0	187	9	US-09-975-719-287
3	60	28.0	270	10	US-09-729-674-132
4	59	27.6	696	9	US-10-122-805-4
5	56.5	26.4	529	12	US-10-042-417-44
6	56	26.2	447	10	US-09-888-615-109
7	55.5	25.9	107	9	US-09-738-626-6443
8	55	25.7	240	9	US-10-042-894A-2
9	54	25.2	1243	9	US-10-196-935A-4
10	53.5	25.0	150	9	US-09-764-891-3210
11	53.5	25.0	150	10	US-09-908-711-123
12	53.5	25.0	217	9	US-10-062-548-113
13	53	24.8	111	10	US-09-864-761-47489
14	53	24.8	205	9	US-10-091-458-38
15	53	24.8	830	9	US-09-870-759-134
16	53	24.8	830	9	US-09-842-758-57
17	52.5	24.5	436	10	US-09-925-300-1640
18	52	24.3	549	10	US-09-764-864-1131
19	52	24.3	830	9	US-09-870-759-140

20	51.5	24.1	478	9	US-09-924-340-108	Sequence 108, App
21	51.5	24.1	478	9	US-09-992-600A-108	Sequence 108, App
22	51.5	24.1	478	9	US-09-746-783-184	Sequence 184, App
23	51.5	24.1	478	9	US-10-000-489-108	Sequence 108, App
24	51.5	24.1	478	9	US-10-000-986-108	Sequence 108, App
25	51	23.8	56	10	US-09-864-761-41214	Sequence 41214, A
26	51	23.8	247	9	US-09-738-626-5773	Sequence 5773, Ap
27	51	23.8	347	10	US-09-789-361-184	Sequence 184, App
28	51	23.8	459	9	US-10-153-668-86	Sequence 86, Appl
29	51	23.8	499	10	US-09-976-165-40	Sequence 40, Appl
30	51	23.8	503	9	US-10-211-962-62	Sequence 62, Appl
31	50.5	23.6	480	9	US-09-796-753-158	Sequence 158, App
32	50.5	23.6	756	10	US-09-946-175-2	Sequence 2, Appli
33	50.5	23.6	761	10	US-09-946-175-3	Sequence 3, Appli
34	50.5	23.6	792	9	US-10-059-585-22	Sequence 22, Appl
35	50.5	23.6	934	9	US-09-796-753-156	Sequence 156, App
36	50.5	23.6	985	9	US-09-978-295A-211	Sequence 211, App
37	50.5	23.6	985	9	US-09-978-697-211	Sequence 211, App
38	50.5	23.6	985	9	US-09-978-192A-211	Sequence 211, App
39	50.5	23.6	985	9	US-09-999-832A-211	Sequence 211, App
40	50.5	23.6	985	9	US-09-978-189-211	Sequence 211, App
41	50.5	23.6	985	9	US-09-978-608A-211	Sequence 211, App
42	50.5	23.6	985	9	US-09-978-191A-211	Sequence 211, App
43	50.5	23.6	985	9	US-09-978-403A-211	Sequence 211, App
44	50.5	23.6	985	9	US-09-978-584A-211	Sequence 211, App
45	50.5	23.6	985	9	US-09-978-585A-211	Sequence 211, App

## ALIGNMENTS

## RESULT 1

US-09-826-581-6  
; Sequence 6, Application US/09826581  
; Patent No. US20020142310A1  
; GENERAL INFORMATION:  
; APPLICANT: Andersson, Leif  
; APPLICANT: Luthman, L. Holger  
; APPLICANT: Marklund, Stefan  
; TITLE OF INVENTION: VARIANTS OF THE HUMAN AMP-ACTIVATED PROTEIN KINASE GAMMA 3 SU  
; FILE REFERENCE: 11145-007001  
; CURRENT APPLICATION NUMBER: US/09/826,581  
; CURRENT FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 60/195,665  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-581-6

Query Match 100.0%; Score 214; DB 10; Length 489;  
Best Local Similarity 100.0%; Pred. No. 4.1e-17;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GKRRAKALRWTRQKSVEGEPPQCGEGPRSRPAAESTGLEA 41

Db 51 GKRRAKALRWTRQKSVEGEPPQCGEGPRSRPAAESTGLEA 91

## RESULT 2

US-09-975-719-287  
; Sequence 287, Application US/09975719  
; Publication No. US20030022349A1  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Rahme, Laurence G.  
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
; FILE REFERENCE: 00786/361003  
; CURRENT APPLICATION NUMBER: US/09/975,719





; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (125)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (138)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (148)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (150)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-3210

Query Match 25.0%; Score 53.5; DB 9; Length 150;  
Best Local Similarity 48.4%; Pred. No. 49;  
Matches 15; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 4 RAKALWTRQKSVEEGEPGPGSRPAA 34  
| : : : | | : | | | | | | | | | | |  
Db 3 RVRSLNTRCRR-EEGEGEGGPGRRRREA 32

RESULT 11  
US-09-908-711-123  
; Sequence 123, Application US/09908711  
; Patent No. US20020045230A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PAL28  
; CURRENT APPLICATION NUMBER: US/09/908,711  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US01/01360  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,867  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01344  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,892  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01345  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,888  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01329  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,905  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01354  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,891  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01339  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,869  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01340  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,874  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01334  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,898  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01320  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,853  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01349  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,902

; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01239  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,870  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01348  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,882  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01347  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,896  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01307  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,864  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01341  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,856  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01336  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,868  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01312  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 167  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 123  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (84)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (125)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (138)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (148)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (150)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-908-711-123

Query Match 25.0%; Score 53.5; DB 10; Length 150;  
Best Local Similarity 48.4%; Pred. No. 49;  
Matches 15; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 4 RAKALWTRQKSVEEGEPGPGSRPAA 34  
| : : : | | : | | | | | | | | | | |  
Db 3 RVRSLNTRCRR-EEGEGEGGPGRRRREA 32

RESULT 12  
US-10-062-548-113  
; Sequence 113, Application US/10062548  
; Publication No. US20030096982A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 44 Human Secreted Proteins

```
; FILE REFERENCE: P2024P1
; CURRENT APPLICATION NUMBER: US/10/062,548
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/369,247
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/074,118
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,157
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,137
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,341
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,141
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (131)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-062-548-113

Query Match      25.0%; Score 53.5; DB 9; Length 217;
Best Local Similarity 48.0%; Pred. No. 72;
Matches 12; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY      14 KSVGEPPQCGEGRSRPAESTG 38
      ::::| | | | | | | | | |
Db      76 RTMDGTGPPAVGE-PRSGPSAGSAG 99

RESULT 13
US-09-864-761-47489
; Sequence 47489, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
```

```
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47489
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009962.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 14
; OTHER INFORMATION: EST HUMAN HIT: AW602053.1, EVALUE 7.00e-25
; OTHER INFORMATION: SWISSPROT HIT: Q16696, EVALUE 3.00e-27
US-09-864-761-47489

Query Match      24.8%; Score 53; DB 10; Length 111;
Best Local Similarity 41.7%; Pred. No. 41;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY      9 RWTQRKSVEEGPPGOGGPRSRP 32
      :| :| :| :| :| :| :| :| :|
Db      65 QWELQSQAEQAGSGEGGRQHRP 88

RESULT 14
US-10-091-458-38
; Sequence 38, Application US/10091458
; Publication No. US20030068627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ06C1
; CURRENT APPLICATION NUMBER: US/10/091,458
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,900
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
```

;  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,597  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509

;  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,581  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17

1	:	PRIOR APPLICATION NUMBER:	60/249,297
2	:	PRIOR FILING DATE:	2000-11-17
3	:	PRIOR APPLICATION NUMBER:	60/232,400
4	:	PRIOR FILING DATE:	2000-09-14
5	:	PRIOR APPLICATION NUMBER:	60/231,242
6	:	PRIOR FILING DATE:	2000-09-08
7	:	PRIOR APPLICATION NUMBER:	60/232,081
8	:	PRIOR FILING DATE:	2000-09-08
9	:	PRIOR APPLICATION NUMBER:	60/232,080
10	:	PRIOR FILING DATE:	2000-09-08
11	:	PRIOR APPLICATION NUMBER:	60/231,414
12	:	PRIOR FILING DATE:	2000-09-08
13	:	PRIOR APPLICATION NUMBER:	60/231,244
14	:	PRIOR FILING DATE:	2000-09-08
15	:	PRIOR APPLICATION NUMBER:	60/233,064
16	:	PRIOR FILING DATE:	2000-09-14
17	:	PRIOR APPLICATION NUMBER:	60/233,063
18	:	PRIOR FILING DATE:	2000-09-14
19	:	PRIOR APPLICATION NUMBER:	60/232,357
20	:	PRIOR FILING DATE:	2000-09-14
21	:	PRIOR APPLICATION NUMBER:	60/232,399
22	:	PRIOR FILING DATE:	2000-09-14
23	:	PRIOR APPLICATION NUMBER:	60/232,401
24	:	PRIOR FILING DATE:	2000-09-14
25	:	PRIOR APPLICATION NUMBER:	60/241,808
26	:	PRIOR FILING DATE:	2000-10-20
27	:	PRIOR APPLICATION NUMBER:	60/241,826
28	:	PRIOR FILING DATE:	2000-10-20
29	:	PRIOR APPLICATION NUMBER:	60/241,786
30	:	PRIOR FILING DATE:	2000-10-20
31	:	PRIOR APPLICATION NUMBER:	60/241,221
32	:	PRIOR FILING DATE:	2000-10-20
33	:	PRIOR APPLICATION NUMBER:	60/246,475
34	:	PRIOR FILING DATE:	2000-11-08
35	:	PRIOR APPLICATION NUMBER:	60/231,243
36	:	PRIOR FILING DATE:	2000-09-08

```

RESULT 15
US-09-870-759-134
; Sequence 134, Application US/09870759
; Patent NO. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-134

```

```
Query Match      24.8%   Score 53; DB 9; Length 830;
Best Local Similarity 44.0%; Pred. NO. 3.3e-02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:55:08 ; Search time 6.60595 Seconds  
(without alignments)  
827.023 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_320\_360

Perfect score: 213

Sequence: 1 HILTHRLKFLHIFGSLP.....PSFLYRTIQDLGICGTRDLA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	181	22	AB111241 Human AMP-activate
2	213	100.0	305	22	AAE00221 Human AMPK gamma s
3	213	100.0	305	22	AAE00328 Human Prkg3 R41Q
4	213	100.0	305	22	AAE00329 Human Prkg3 V40I
5	213	100.0	464	22	AAE00223 Human AMPK gamma s
6	213	100.0	489	22	AAE00223 PRKAG3. Homo sapi
7	210	98.6	305	22	AAE00220 Pig AMPK gamma sub
8	210	98.6	305	22	AAE00225 Sus scrofa AMPK ga
9	210	98.6	305	22	AAE00226 Sus scrofa AMPK ga
10	210	98.6	464	22	AAE00226 Pig AMPK gamma sub

11	210	98.6	464	23	AAE22984 Pig wild-type PRKA
12	210	98.6	464	23	AAE22985 Pig PRKAG3 polymor
13	210	98.6	464	23	AAE22986 Pig PRKAG3 polymor
14	210	98.6	464	23	AAE22987 Pig PRKAG3 polymor
15	210	98.6	464	23	AAE22988 Pig PRKAG3 polymor
16	210	98.6	514	22	AAE00224 Sus scrofa Prkg3
17	124	58.2	328	20	AAW88438 Disease associated
18	124	58.2	328	20	AAW88432 Human protein sequ
19	124	58.2	383	22	ABG20080 Novel human diago
20	124	58.2	488	22	ABG20078 Novel human diago
21	112	52.6	331	18	AAW29817 Mammalian AMPK-gam
22	112	52.6	344	21	AAE54009 Human pancreatic c
23	112	52.6	353	23	ABB06101 Human NS protein s
24	105	49.3	634	22	ABB66245 Drosophila melanog
25	84	39.4	180	22	AAW75798 Human colon cancer
26	62.5	29.3	318	21	AAW67888 Soybean sucrose no
27	59	27.7	105	23	ABB89679 Human polypeptide
28	55	25.8	798	19	AAW33751 Thermostable phosp
29	54.5	25.6	172	11	AAW05412 Part of human alph
30	54.5	25.6	452	10	AAW90534 Peptide sequence o
31	54.5	25.6	464	11	AAW05411 Pro-type human pla
32	54.5	25.6	488	21	AAW59191 Human alpha-2 anti
33	54.5	25.6	490	11	AAW04252 Amino acid sequenc
34	54.5	25.6	491	10	AAW93423 Human alpha-2 plas
35	54.5	25.6	491	12	AAW13860 Human alpha-2 plas
36	54.5	25.6	491	21	AAW59192 Bovine alpha-2 ant
37	54.5	25.6	492	10	AAW90466 Human alpha-2 plas
38	54.5	25.6	744	9	AAW81006 Alpha-2-plasmin in
39	54	25.4	92	20	AAW13213 Human secreted pro
40	54	25.4	263	22	AAW87756 Human T2R26 amino
41	54	25.4	1156	22	ABG09141 Novel human diago
42	54	25.4	1156	22	ABG18480 Novel human diago
43	53	24.9	79	22	ABG22403 Novel human diago
44	53	24.9	357	22	AAW87274 Novel central nerv
45	53	24.9	684	22	ABG22406 Novel human diago

ALIGNMENTS

RESULT 1	
AB111241	
ID	AB111241 standard; peptide: 181 AA.
XX	AB111241:
AC	
XX	
XX	
DT	11-JAN-2002 (first entry)
XX	
DE	Human AMP-activated protein kinase subunit homologue, SEQ ID NO:1611.
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW	inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW	chronic inflammatory condition; proliferative retinopathy;
KW	atherosclerosis; coronary heart disease; arterial ischaemia;
KW	bone disorder; osteoporosis; vascular growth disorder;
KW	tissue regeneration; wound healing; infection; immune disorder;
KW	cell culture; drug screening; gene therapy; antiinflammatory;
KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW	antifungal; vulnerary; antiulcer.
XX	
OS	Homo sapiens.
XX	
XX	
PN	W0200157188-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US03800.
XX	
PR	03-FEB-2000; 2000US-0496914.



PR 27-APR-2000; 2000US-0560675.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

DR N-PSDB; ABA08485.

XX Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -

XX Claim 20; Page 159-160; 1963pp; English.

XX Sequences ABA10981-AB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis; cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention.

XX Sequence 181 AA;

XX Query Match 100.0%; Score 213; DB 22; Length 181;  
XX Best Local Similarity 100.0%; Pred. No. 1.6e-23;

XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHFGSLLPSPFLYRTIQDLGIGTFRDLA 41

DB 80 HILTHKLLKFLHFGSLLPSPFLYRTIQDLGIGTFRDLA 120

RESULT 2

AAE00221

ID AAE00221 standard; Protein; 305 AA.

XX AAE00221;

XX AAE00221;

XX 13-JUN-2001 (first entry)

XX Human AMPK gamma subunit muscle-specific isoform, PRKAG3.

DE

XX

KW Human: gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy.

XX Homo sapiens.

OS WO200120003-A2.

PN 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP09896.

XX 10-SEP-1999; 99EP-0402236.

XX 18-MAY-2000; 2000EP-0401388.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX (ANDE/) ANDERSSON L.

XX (LOOF/) LOOFT C.

XX (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

XX Tannuccelli N, Gellin J, Le Roy P, Chardon P;

XX WPI; 2001-244810/25.

XX N-PSDB; AAD03296.

XX New variants of the gamma subunit of vertebrate adenosine

XX monophosphate-activated kinase for diagnosis or treatment of disorders

XX associated with energy metabolism such as diabetes, obesity, and

XX myopathy -

XX Claim 4; Page 55-57; 71pp; English.

XX The present sequence is human adenosine monophosphate

XX (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,

XX PRKAG3. Mutation in Prkag3 results in an altered regulation of

XX carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is

XX useful as therapeutic for treating carbohydrate metabolism disorders such

XX as diabetes, obesity, and disorders associated with muscle metabolism

XX such as myopathy and cardiovascular diseases, to modulate AMPK

XX activity, and for restoring a normal AMPK function. PRKAG3 sequence

XX and its functionally altered mutants are useful for the diagnostic

XX evaluation, genetic testing and prognosis of a metabolic disorder,

XX preferably a carbohydrate metabolism disorder. Primers that can detect

XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

XX useful for detecting a dysfunction of carbohydrate metabolism resulting

XX from the expression of a functionally altered allele of PRKAG3.

XX Transgenic animal and host cell transformed with PRKAG3 or a

XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

XX screening compounds able to modulate AMPK activity. Nucleic acid

XX encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or

XX in a sequence encoding the first cystathione beta synthase (CBS) domain

XX of PRKAG3 and is useful in gene therapy.

XX Sequence 305 AA;

XX Query Match 100.0%; Score 213; DB 22; Length 305;

XX Best Local Similarity 100.0%; Pred. No. 2.8e-23;

XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHFGSLLPSPFLYRTIQDLGIGTFRDLA 41

DB 136 HILTHKLLKFLHFGSLLPSPFLYRTIQDLGIGTFRDLA 176

RESULT 3

AAE00328

ID AAE00328 standard; Protein; 305 AA.

XX AAE00328;

XX AAE00328;

XX

DT 13-JUN-2001 (first entry)  
XX Human Prkg3 R41Q mutant.  
DE  
XX  
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 41 /note= "Wild-type Arg substituted with Gln"  
FT  
XX  
XX  
PN WO200120003-A2.  
XX  
XX  
PD 22-MAR-2001.  
XX  
XX  
PF 11-SEP-2000; 2000WO-EP09896.  
XX  
XX 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.  
PR  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
PI  
XX WPI; 2001-244810/25.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
XX monophosphate-activated kinase for diagnosis or treatment of disorders  
XX associated with energy metabolism such as diabetes, obesity, and  
XX myopathy .  
XX  
XX Disclosure; Page -: 71pp; English.  
XX  
XX The present sequence is a R41Q mutant of human muscle-specific isoform  
XX of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
XX (AMPK) Prkg3. This mutant sequence results in increased glycogen  
XX content in human skeletal muscle.  
XX Mutation in Prkg3 results in an altered regulation of carbohydrate  
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
XX therapeutic for treating carbohydrate metabolism disorders such as  
XX diabetes, obesity, and disorders associated with muscle metabolism  
XX such as myopathy and cardiovascular diseases, to modulate AMPK  
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence  
XX and its functionally altered mutants are useful for the diagnostic  
XX evaluation, genetic testing and prognosis of a metabolic disorder,  
XX preferably a carbohydrate metabolism disorder. Primers that can detect  
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
XX useful for detecting a dysfunction of carbohydrate metabolism resulting  
XX from the expression of a functionally altered allele of PRKAG3.  
XX Transgenic animal and host cell transformed with PRKAG3 or a  
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
XX screening compounds able to modulate AMPK activity. Nucleic acid  
XX encoding PRKAG3 is useful for detecting mutations in a Prkg3 gene, or  
XX in a sequence encoding the first cystathione beta synthase (CBS) domain  
XX of PRKAG3 and is useful in gene therapy.  
XX Note: The present sequence is not shown in the specification, but  
XX is derived from the human Prkg3 sequence SEQ.ID.NO.4 shown in page 57-58  
XX of sequence listing (AAE00221).  
XX  
SQ Sequence 305 AA;

Query Match 100.0%; Score 213; DB 22; Length 305;  
Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKLFHFGSLPRPSFLYRTIQLGIGCTFRDLA 41  
XX  
DB 136 HILTHKRLKLFHFGSLPRPSFLYRTIQLGIGCTFRDLA 176  
RESULT 4  
AAE00329  
ID AAE00329 standard; Protein; 305 AA.  
XX  
XX AC AAE00329;  
XX  
XX 13-JUN-2001 (first entry)  
XX  
XX Human Prkg3 V40I mutant.  
XX  
XX Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 40 /note= "Wild-type Val substituted with Ile"  
FT  
XX  
XX WO200120003-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 11-SEP-2000; 2000WO-EP09896.  
XX  
XX 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.  
XX  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
PI  
XX WPI; 2001-244810/25.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
XX monophosphate-activated kinase for diagnosis or treatment of disorders  
XX associated with energy metabolism such as diabetes, obesity, and  
XX myopathy .  
XX  
XX Disclosure; Page -: 71pp; English.  
XX  
XX The present sequence is a V40I mutant of human muscle-specific isoform  
XX of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
XX (AMPK) Prkg3. This mutant sequence results in decreased glycogen  
XX content in human skeletal muscle.  
XX Mutation in Prkg3 results in an altered regulation of carbohydrate  
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
XX therapeutic for treating carbohydrate metabolism disorders such as  
XX diabetes, obesity, and disorders associated with muscle metabolism  
XX such as myopathy and cardiovascular diseases, to modulate AMPK  
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence  
XX and its functionally altered mutants are useful for the diagnostic  
XX evaluation, genetic testing and prognosis of a metabolic disorder,  
XX preferably a carbohydrate metabolism disorder. Primers that can detect  
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
XX useful for detecting a dysfunction of carbohydrate metabolism resulting  
XX from the expression of a functionally altered allele of PRKAG3.  
XX Transgenic animal and host cell transformed with PRKAG3 or a  
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
XX screening compounds able to modulate AMPK activity. Nucleic acid  
XX encoding PRKAG3 is useful for detecting mutations in a Prkg3 gene, or  
XX in a sequence encoding the first cystathione beta synthase (CBS) domain  
XX of PRKAG3 and is useful in gene therapy.  
XX Note: The present sequence is not shown in the specification, but  
XX is derived from the human Prkg3 sequence SEQ.ID.NO.4 shown in page 57-58  
XX of sequence listing (AAE00221).  
XX  
SQ Sequence 305 AA;

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the human Prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
 CC of sequence listing (AAE00221).

XX Sequence 305 AA;  
 SQ Query Match 100.0%; Score 213; DB 22; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 Db 136 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 176

RESULT 5  
 AAE00223  
 ID AAE00223 standard; Protein; 464 AA.  
 XX  
 AC AAE00223;  
 XX  
 DF 13-JUN-2001 (first entry)  
 DT  
 DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
 XX  
 KW Human; gamma subunit; adenosine monophosphate-activated Kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiand; gene therapy.

XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Domain 172..225  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Misc-difference 200  
 FT /note= "RN- mutation site"  
 FT Domain 253..307  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 329..382  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 400..453  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"

XX WO200120003-A2.  
 PN  
 XX  
 XX 22-MAR-2001.  
 PD  
 XX  
 XX 11-SEP-2000; 2000WO-EP09896.  
 PF  
 XX  
 XX 10-SEP-1999; 99EP-0402236.  
 PR  
 XX 18-MAY-2000; 2000EP-0401388.  
 PR  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX  
 XX WPI; 2001-244810/25.  
 DR N-PSDB; AAD03320.

XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and  
 XX myopathy -  
 XX Claim 5; Fig 3; 71pp; English.  
 XX  
 CC The present sequence is human adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform, of  
 CC complete PRKAG3. Mutation in Prkag3 results in an altered regulation of  
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
 CC useful as therapeutic for treating carbohydrate metabolism disorders such  
 CC as diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.

XX Sequence 464 AA;

Query Match 100.0%; Score 213; DB 22; Length 464;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 Db 295 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 6  
 AAB47679  
 ID AAB47679 standard; Protein; 489 AA.  
 XX  
 AC AAB47679;  
 XX  
 XX 21-JAN-2002 (first entry)  
 DT  
 DE PRKAG3.  
 DE  
 XX  
 XX Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;  
 KW metabolic disease; diabetes; obesity; substitution; ss.  
 KW Homo sapiens.

OS  
 XX Key Location/Qualifiers  
 XX Misc-difference 71  
 FT /note= "Possible variation point P71A"  
 FT Misc-difference 340  
 FT /note= "Possible variation point R340W"  
 FT  
 PN WO200177305-A2.

XX 18-OCT-2001.  
 PD  
 XX 06-APR-2001; 2001WO-SE00765.  
 XX  
 PF  
 XX 07-APR-2000; 2000US-195665P.  
 PR  
 XX (AREX-) AREXIS AB.  
 XX  
 PA Andersson L, Luthman H, Marklund S;

PI  
 XX WPI; 2001-657170/75.  
 DR N-PSDB; AAH43685.

XX

PT New variants of human AMP-activated protein kinase gamma3 subunit  
 PT associated with a metabolic disease e.g. diabetes or obesity and method  
 PT for determining a risk estimate of diseases in subject by detecting the  
 PT variant -

XX Disclosure: Fig 5; 25pp; English.

XX This sequence is encoded by the full length cDNA encoding the human  
 CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting  
 CC the presence of the PRKAG3 DNA, or a variant, is useful in determining  
 CC a risk estimate of a metabolic disease, such as diabetes or obesity,  
 CC in a subject. The variation may occur in exons 3, 4 or 10. In exon  
 CC 3 variation may be a substitution of a G for a C at nucleotide 320,  
 CC resulting in the amino acid substitution P71A; in exon 4 variation may  
 CC be a substitution of a T for a C at nucleotide 550; and in exon 10  
 CC variation may be a substitution of a T for a C at nucleotide 1037,  
 CC resulting in the amino acid substitution R349W. There may also be  
 CC nucleotide variation in intron 6. The numbering of these  
 CC variations is based on the full length cDNA, rather than on  
 CC position 1 of the open reading frame.

XX Sequence 489 AA;

Query Match 100.0%; Score 213; DB 22; Length 489;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 41  
 |||||||||||||||||||||||||||||||||||||||||  
 Db 320 HILTHKLLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 360

RESULT 7

AAE00220  
 ID AAE00220 standard; Protein; 305 AA.

AC AAE00220;

DT 13-JUN-2001 (first entry)

DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3.

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;  
 KW Chromosome 15.

XX Sus scrofa.

XX Key Location/Qualifiers

FT Domain 13..66  
 /label= CBS  
 /note= "Cystathione beta synthase domain"  
 FT Domain 94..148  
 /label= CBS  
 /note= "Cystathione beta synthase domain"  
 FT Domain 170..223  
 /label= CBS  
 /note= "Cystathione beta synthase domain"  
 FT Domain 241..294  
 /label= CBS  
 /note= "Cystathione beta synthase domain"

PN WC0200120003-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP09896.

XX 10-SEP-1999; 99EP-0402236.

XX 18-MAY-2000; 2000EP-0401388.

PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Tannuccelli N, Gellin J, Le Roy P, Chardon P;  
 DR WPI; 2001-244810/25.  
 DR N-PSDB; AAD03295.

XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -

XX Claim 4; Fig 2; 71pp; English.

XX The present amino acid sequence is pig adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
 CC PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.  
 CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.

XX Sequence 305 AA;

Query Match 98.6%; Score 210; DB 22; Length 305;  
 Best Local Similarity 97.6%; Pred. No. 7.8e-23;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 41  
 |||||||||||||||||||||||||||||||||||||||||  
 Db 136 HILTHKLLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 176

RESULT 8

AAE00225  
 ID AAE00225 standard; Protein; 305 AA.

AC AAE00225;

DT 13-JUN-2001 (first entry)

XX Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).

DE Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
 KW variant.

XX Sus scrofa.

XX Key Location/Qualifiers

FT Domain 13..66

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

```

FT Misc-difference 41 /note= "Wild-type Arg substituted with Gln"
FT Domain 94..148
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
PN W0200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
XX WPI; 2001-244810/25.
XX
XX New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 9; Page -: 7lpp; English.
XX
XX The present sequence is a R41Q mutant of muscle-specific isoform of
CC gamma subunit of adenosine monophosphate (AMP)-activated kinase
CC (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in
CC increased glycogen content in pig skeletal muscle.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
CC Note: The present sequence is not shown in the specification, but
CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).
XX
SQ Sequence 305 AA;

```

Query Match 98.6%; Score 210; HB 22; Length 305;  
Best Local Similarity 97.6%; Pred. No. 7.Re-23;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLIPRPSFLYRTIQLDGICTFRDLA 41  
|||||  
DB 136 HILTHKRLKFLHIFGSLIPRPSFLYRTIQLDGICTFRDLA 176

```

RESULT 9
AAE00226
ID AAE00226 standard; Protein; 305 AA.
XX
XX
AC AAE00226;
XX
XX 13-JUN-2001 (first entry)
XX
XX Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).
XX
XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; muten;
KW variant.
XX
XX Sus scrofa.
XX
XX Key Location/Qualifiers
FH Domain 13..66
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Misc-difference 40
FT /note= "Wild-type Val substituted with Ile"
FT Domain 94..148
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
XX W0200120003-A2.
XX
XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-EP09896.
XX
XX 10-SEP-1999; 99EP-0402236.
XX
XX 18-MAY-2000; 2000EP-0401388.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFT C.
XX (KALM/) KALM E.
XX
XX Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
XX WPI; 2001-244810/25.
XX
XX New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 9; Page -: 7lpp; English.
XX
XX The present sequence is a V40I mutant of muscle-specific isoform of
CC gamma subunit of adenosine monophosphate (AMP)-activated kinase
CC (AMPK) prkag3 from Sus scrofa. This mutant sequence results in
CC decreased glycogen content in pig skeletal muscle.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
CC Note: The present sequence is not shown in the specification, but
CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).
XX
SQ Sequence 305 AA;

```

CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).  
 XX  
 SQ Sequence 305 AA;  
 Query Match 98.6%; Score 210; DB 22; Length 305;  
 Best Local Similarity 97.6%; Pred. No. 7.8e-23;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 Db 136 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 176  
 RESULT 10  
 AAE00222  
 ID AAE00222 standard; Protein; 464 AA.  
 AC AAE00222;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
 KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;  
 KW chromosome 15.  
 OS  
 XS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 172..225  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Misc-difference 200  
 FT /note= "RN- mutation site"  
 FT Domain 253..307  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 329..392  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 400..453  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 XX  
 PN WO200120003-A2.  
 XX  
 PD 22-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-EP09896.  
 XX  
 PR 10-SEP-1999; 99EP-0402236.  
 PR 18-MAY-2000; 2000EP-0401388.  
 XX  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX  
 DR WPI; 2001-244810/25.  
 DR N-PSDB; AAD03319.  
 XX  
 FT New variants of the gamma subunit of vertebrate adenosine  
 FT monophosphate-activated kinase for diagnosis or treatment of disorders  
 FT associated with energy metabolism such as diabetes, obesity, and  
 FT myopathy  
 XX  
 PS Claim 5; Fig 3: 7lpp; English.  
 XX  
 CC The present sequence is pig adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
 CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome  
 CC 15. Mutation in prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 98.6%; Score 210; DB 22; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 1.3e-22;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 Db 295 HILTHKRLKFLHIFGTLPRPSFLYRTIQDLGIGTFRDLA 335  
 RESULT 11  
 AAE22984  
 ID AAE22984 standard; Protein; 464 AA.  
 XX  
 AC AAE22984;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig wild-type PRKAG3 protein.  
 XX  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 30  
 FT /note= "Wild type Asn is replaced with Thr during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 52  
 FT /note= "Wild type Gly is replaced with Ser during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 199  
 FT /note= "Wild type Val is replaced with Ile during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 200  
 FT /note= "Wild type Arg is replaced with Gln during  
 FT single nucleotide polymorphism (SNP)"  
 FT



```

XX WPI: 2002-393850/42.
DR N-PSDB; AAD36458.
XX
XX Screening animals to determine those likely to produce larger litters
PT and improved meat quality traits involves assaying for the presence of
PT polymorphisms in the AMP activated protein kinase regulatory gamma
PT subunit gene -
XX
XX Claim 36; Page 96-97; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype
CC in the sample of genetic material obtained from animal. The genotype
CC is characterised by polymorphism(s) in the AMP activated protein
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used
CC for screening animals e.g., pigs to determine those most likely to
CC exhibit improved meat quality traits and to produce larger litters.
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX Sequence 464 AA;
SQ
Query Match 98.6%; Score 210; DB 23; Length 464;
Best Local Similarity 97.6%; Pred. No. 1.3e-22;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB |||||||||||||:|||||||||||||||||||||
295 HILTHKRLKFLHFGTLLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 14
AAE22987
ID AAE22987 standard; Protein; 464 AA.
XX
XX AAE22987;
AC
XX
XX 09-AUG-2002 (first entry)
DT
XX
XX Pig PRKAG3 polymorphic variant (PRKAG3-199).
DE
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
KW screening; meat quality; single nucleotide polymorphism; SNP; pig;
KW variant.
XX
XX Sus scrofa.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 199
FT /note= "Wild type Val is substituted with Ile due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 10-SEP-2001; 2001WO-US28283.
PF
XX
XX 08-SEP-2000; 2000US-231045P.
PR
XX 08-JAN-2001; 2001US-260239P.
PR
XX 18-JUN-2001; 2001US-299111P.
PR
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
PA
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
PI
XX WPI: 2002-393850/42.
DR N-PSDB; AAD36459.
DR
XX
XX Screening animals to determine those likely to produce larger litters
PT and improved meat quality traits involves assaying for the presence of
PT polymorphisms in the AMP activated protein kinase regulatory gamma
PT subunit gene -
XX
XX Disclosure; Page 105-107; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype

```

```

PT subunit gene -
XX
XX Disclosure; Page 100-102; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype
CC in the sample of genetic material obtained from animal. The genotype
CC is characterised by polymorphism(s) in the AMP activated protein
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used
CC for screening animals e.g., pigs to determine those most likely to
CC exhibit improved meat quality traits and to produce larger litters.
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-199).
XX
XX Sequence 464 AA;
SQ
Query Match 98.6%; Score 210; DB 23; Length 464;
Best Local Similarity 97.6%; Pred. No. 1.3e-22;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB |||||||||||||:|||||||||||||||||||||
295 HILTHKRLKFLHFGTLLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 15
AAE22988
ID AAE22988 standard; Protein; 464 AA.
XX
XX AAE22988;
AC
XX
XX 09-AUG-2002 (first entry)
DT
XX
XX Pig PRKAG3 polymorphic variant (PRKAG3-200).
DE
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
KW screening; meat quality; single nucleotide polymorphism; SNP; pig;
KW variant.
XX
XX Sus scrofa.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 200
FT /note= "Wild type Arg is substituted with Gln due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 10-SEP-2001; 2001WO-US28283.
PF
XX
XX 08-SEP-2000; 2000US-231045P.
PR
XX 08-JAN-2001; 2001US-260239P.
PR
XX 18-JUN-2001; 2001US-299111P.
PR
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
PA
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
PI
XX WPI: 2002-393850/42.
DR N-PSDB; AAD36460.
DR
XX
XX Screening animals to determine those likely to produce larger litters
PT and improved meat quality traits involves assaying for the presence of
PT polymorphisms in the AMP activated protein kinase regulatory gamma
PT subunit gene -
XX
XX Disclosure; Page 105-107; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
CC those more likely to produce large litters and improved meat quality
CC traits. The method involves assaying for the presence of a genotype

```



CC in the sample of genetic material obtained from animal. The genotype  
CC is characterised by polymorphism(s) in the AMP activated protein  
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
CC for screening animals e.g., pigs to determine those most likely to  
CC exhibit improved meat quality traits and to produce larger litters.  
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).  
XX

SQ Sequence 464 AA;

Query Match 98.6%; Score 210; DB 23; Length 464;  
Best Local Similarity 97.6%; Pred. No. 1.3e-22;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
|||||  
DB 295 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 335

Search completed: June 6, 2003, 11:01:03  
Job time : 7.60595 secs

Result No.	Query	Score	Match	Length	DB	ID	Description
1	112	52.6	330	2	T10759	AMP-activated prote	
2	97	45.5	478	2	T25899	hypothetical prote	
3	63	29.6	379	2	T10971	p442 protein - kid	
4	56	26.3	1270	2	T26720	hypothetical prote	
5	55.5	26.1	516	2	H21145	conserved hypothet	
6	54.5	25.6	491	1	ITHU42	alpha-2-antiplasmi	
7	54	25.4	269	2	T39029	hypothetical prote	
8	53	24.9	305	1	S59244	hypothetical prote	
9	53	24.9	305	1	S67067	coproporphyrinogen	
10	52.5	24.6	492	2	S67067	probable membrane	
11	52	24.4	557	2	S43977	alpha-2-antiplasmi	
12	51.5	24.2	371	2	S61980	histone acetyltran	
13	51.5	24.2	527	2	E87536	ABC transporter, p	
14	51.5	24.2	629	2	T18227	hypothetical prote	
15	51	23.9	205	2	C64677	conserved hypothet	
16	51	23.9	234	2	A71838	hypothetical prote	
17	51	23.9	309	2	A90368	coproporphyrinogen	
18	51	23.9	373	2	D90032	hypothetical prote	
19	51	23.9	564	2	F97601	afub (AE006182) [i	
20	51	23.9	564	2	AE2823	ABC transporter, m	
21	51	23.9	583	2	D82634	acetolactate synth	
22	50.5	23.7	205	2	AB1903	hypothetical prote	
23	50.5	23.7	295	2	H82600	hypothetical prote	
24	50.5	23.7	412	2	S76239	biotin synthesis p	
25	50.5	23.7	470	2	F83202	hypothetical prote	
26	50	23.5	160	2	B81132	probable phosphogl	
27	50	23.5	160	2	D81892	conserved hypothet	
28	50	23.5	181	2	T36787	hypothetical prote	
29	50	23.5	1402	2	S75938	probable NPP pyrop	
30	50	23.5	1402	2	S75938	chemotaxis protein	

```

Best Local Similarity 48.6%; Pred. No. 11;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 4 THKRLKLFHFGSLPRPSFLYRTIQDLGIGTFR 38
      | : | : | | | | | | | | | | | |
Db 128 TLQRQQVLHA-GSGPCLPHLLSRLCQDLGPGAFR 161

RESULT 7
T39029
hypothetical protein SPAC6C3.05 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39029
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsby, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z21750
A:Accession: T39029
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-269 <DEV>
A:CROSS-references: EMBL:Z69731; PIDN:CAB40280.1; GSPDB:GN00066; SPDB:SPAC6C3.05
A:Experimental source: strain 972h-; cosmid c6C3
C:Genetics:
A:Gene: SPDB:SPAC6C3.05
A:Map position: 1

Query Match 25.4%; Score 54; DB 2; Length 269;
Best Local Similarity 35.9%; Pred. No. 6.9;
Matches 14; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 2 ILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTFRDL 40
      | | | : | | | | | | | | | |
Db 216 IKTHYKTLRIHTFLTFVPVSLSSNLKKQMGWFDL 254

RESULT 8
S52924
coproporphyrinogen oxidase (EC 1.3.3.3) oxygen-dependent - Pseudomonas aeruginosa
N:Alternate names: coproporphyrinogen III oxidase, aerobic
C:Species: Pseudomonas aeruginosa
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Dec-2000
C:Accession: S52924; F83641
R:Hungerer, C.; Troup, B.; Jahn, D.
submitted to the EMBL Data Library, February 1995
A:Description: Cloning and regulation of the Pseudomonas aeruginosa hemF gene encoding
A:Reference number: S52923
A:Accession: S52924
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <HUN>
A:CROSS-references: EMBL:X85015; NID:g747872; PIDN:CAA59376.1; PID:g695693
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83641
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <STO>
A:CROSS-references: GB:AE004442; GB:AE004091; NID:g9945843; PIDN:AGC03414.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: hemF; PA0024
A:Start codon: GTG
C:Superfamily: coproporphyrinogen oxidase
C:Keywords: oxidoreductase

Query Match 24.9%; Score 53; DB 1; Length 305;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 14; Conservative 3; Mismatches 15; Indels 6; Gaps 1;

```

```

QY      3 LTHKRLKFLHIFGSLLP-----RPSFLYRTIQDLGI 34
      | | : | : | | | | | | | | | | | | | | | |
Db      55 LIEKGWNTSHVFGDSLPPSASAHRPGLAGRGFQALGV 92

RESULT 9
S67067
probable membrane protein YOR175c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3635
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67067; S67063
R:Hughes, B.; Pohl, T.M.
A:Reference number: S66685
A:Accession: S67067
A:Molecule type: DNA
A:Residues: 1-619 <HUG>
A:Cross-references: EMBL:Z75083; NID:gl420424; PID:e252056; PID:gl420425; MIPS:YOR175c
A:Experimental source: strain S288C
R:Bordonne, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarassov, I.A.; Winsor
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67032
A:Accession: S67063
A:Molecule type: DNA
A:Residues: 270-619 <BOR>
A:Cross-references: EMBL:Z75083; MIPS:YOR175c
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005701
A:Map position: 15R
C:Keywords: transmembrane protein
F:53-69/Domain: transmembrane #status predicted <TM1>
F:461-477/Domain: transmembrane #status predicted <TM2>

Query Match      24.9%; Score 53; DB 2; Length 619;
Best Local Similarity 42.5%; Pred. No. 24;
Matches 17; Conservative 2; Mismatches 17; Indels 4; Gaps 2;

QY      5 HKRLKFL-----HFGSLPRPSFLYRTTQD-LGIGTFRDL 40
      | | | | | | | | | | | | | | | | | | | | | |
Db      170 HPPLKFLAYAFYFTLLTGPSFDYADFDSWLNCMFRLD 209

RESULT 10
S43977
alpha-2-antiplasmin precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Oct-1994 #sequence_revision 19-May-1995 #text_change 16-Jul-1999
C:Accession: S43977; S27260
R:Christensen, S.; Berglund, L.; Sottrup-Jensen, L.
FEBS Lett. 343, 223-228, 1994
A:Title: Primary structure of bovine alpha-2-antiplasmin.
A:Reference number: S43977; MUID:94229242; PMID:7513654
A:Accession: S43977
A:Molecule type: mRNA
A:Residues: 1-492 <CHR>
A:Cross-references: GB:X78436; NID:g498821; PIDN:CAA55200.1; PID:g498822
A:Experimental source: liver
R:Christensen, S.; Sottrup-Jensen, L.
FEBS Lett. 312, 100-104, 1992
A:Title: Bovine alpha(2)-antiplasmin. N-terminal and reactive site sequence.
A:Reference number: S27260; MUID:93050153; PMID:1385210
A:Accession: S27260
A:Molecule type: protein
A:Residues: 23-27,'Q',29-39,'P',41-42,'E',44-45:374-415 <CH2>
C:Superfamily: antithrombin III
C:Keywords: glycoprotein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-492/product: alpha-2-antiplasmin #status predicted <MAT>
F:71-144/Disulfide bonds: #status experimental
F:127,249,296,310,317/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

```

Query Match      24.6%; Score 52.5; DB 2; Length 492;
Best Local Similarity 47.2%; Pred. No. 22;
Matches 17; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

QY      4 THKRLKFLHI-FGSLPRPSFLYRTTQDLGIGTF 38
      | | : | : | | | | | | | | | | | | | | | |
Db      129 TLQRLKEVLHADSGPCLPH--LLSLRLCDLGGGA 162

RESULT 11
S61980
histone acetyltransferase (EC 2.3.1.48), RNA polymerase II-associated [validated] -
N:Alternate names: protein Lpg22c; protein YPL086c
C:Species: Saccharomyces cerevisiae
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 19-Apr-2002
C:Accession: S61980
R:Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.;
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A:Reference number: S61959
A:Accession: S61980
A:Molecule type: DNA
A:Residues: 1-557 <WAN>
A:Cross-references: EMBL:U43281; NID:gil151218; PIDN:AAB68213.1; PID:gil151240; GSPDB:
C:Genetics:
A:Gene: SGD:ELP3; ELP3; MIPS:YPL086c
A:Cross-references: SGD:S0006007
A:Map position: 16L
C:Function:
A:Description: EC 2.3.1.48 [validated, MUID:99374060]; acetylates core histones in vi
ed in chromatin
C:Superfamily: hypothetical protein YPL086c
C:Keywords: acyltransferase; coenzyme A; transcription

Query Match      24.4%; Score 52; DB 2; Length 557;
Best Local Similarity 37.9%; Pred. No. 29;
Matches 11; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY      8 LKFLHIFGSLPRPSFLYRTTQDLGIGT 36
      | | : | : | | | | | | | | | | | | | | | |
Db      482 IVRELHVGVSVPLHSRDPKFKHQHGGT 510

RESULT 12
E87536
ABC transporter, permease protein, probable CC2318 [imported] - Caulobacter crescenti
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87536
R:Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Winn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87536
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STO>
A:Cross-references: GB:AE005673; NID:gil3423841; PIDN:AAK24289.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2318

Query Match      24.2%; Score 51.5; DB 2; Length 371;
Best Local Similarity 25.6%; Pred. No. 22;
Matches 11; Conservative 11; Mismatches 14; Indels 7; Gaps 1;

QY      2 ILTHKRLKFLHIFGSL-----RPSFLYRTIQDLGIGTF 37
      | | : | : | | | | | | | | | | | | | | | |
Db      77 IIAKQVLRLLLELGAIVKVEPQRAARTGFAALLERIGRVF 119

```

C64677  
 Conserved hypothetical protein HPL259 - *Helicobacter pylori* (strain 28695)  
 C:Species: *Helicobacter pylori*  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 11-Jan-2000  
 C:Accession: C64677  
 R:Tombs, J.F.; White, O.; Kervagade, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 1.50788 Seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_320\_360

Perfect score: 213

Sequence: 1 HILTHKRLKFLHFGSLP.....PSFLYRTIQDLGIGTRDLA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	213	100.0	464	1 AAKI_HUMAN	Q9ug19 homo sapien
2	210	98.6	464	1 AAKI_PIG	Q9my94 sus scrofa
3	124	58.2	569	1 AAKH_HUMAN	Q9ug10 homo sapien
4	112	52.6	330	1 AAKG_BOVIN	P58108 bos taurus
5	112	52.6	330	1 AAKG_MOUSE	O54950 mus musculus
6	112	52.6	330	1 AAKG_RAT	P80355 rattus norv
7	112	52.6	331	1 AAKG_HUMAN	P54619 homo sapien
8	64.5	30.3	133	1 AAKG_PIG	Q09138 sus scrofa
9	54.5	25.6	491	1 A2AP_HUMAN	P08697 homo sapien
10	54	25.4	269	1 YD55_SCHPO	Q10308 schizosacch
11	53	24.9	305	1 HEM6_PSEAE	P43898 pseudomonas
12	53	24.9	1024	1 SZ6L_HUMAN	Q2bvh1 homo sapien
13	52.5	24.6	492	1 A2AP_BOVIN	P28800 bos taurus
14	51.5	24.2	527	1 YB64_YEAST	P38314 saccharomyc
15	51	23.9	309	1 HEM6_YERPE	Q9zc19 yersinia pe
16	50	23.5	1167	1 ITAE_MOUSE	Q60677 mus musculus
17	49.5	23.2	606	1 ABD4_MOUSE	O89016 mus musculus
18	49.5	23.2	2136	1 YCF2_MARPO	P09975 marchantia
19	49	23.0	304	1 LST_HAEIN	Q48211 haemophilus
20	49	23.0	712	1 IRA1_HUMAN	P51617 homo sapien
21	48.5	22.8	1786	1 UVR4_CHLTR	O84337 chlamydia t
22	48.5	22.8	1840	1 SUIS_RAT	P23739 rattus norv
23	48	22.5	186	1 Y786_METJA	Q58196 methanococc
24	48	22.5	210	1 YD07_HAEIN	Q57320 haemophilus
25	48	22.5	320	1 MTDL_YEAST	Q20246 saccharomyc
26	48	22.5	420	1 YAME_RHISN	P55564 rhizobium s
27	47.5	22.3	494	1 CATA_PEA	P25890 pisum sativ
28	47	22.1	271	1 LCR6_YERPE	P28808 yersinia pe
29	47	22.1	271	1 VIRE_YEREN	P13225 yersinia en
30	47	22.1	454	1 ATTY_RAT	P04694 rattus norv
31	47	22.1	798	1 CDB2_HUMAN	Q9y5e7 homo sapien
32	46.5	21.8	161	1 CRAA_TRLIN	P02500 trichechus
33	46.5	21.8	353	1 BRB1_HUMAN	P46663 homo sapien

34	46.5	21.8	606	1 ABD4_HUMAN	O14678 homo sapien
35	46	21.6	184	1 RR4_PATFR	O20273 paterosonia
36	46	21.6	196	1 RR4_PATSQ	O36052 paterosonia
37	46	21.6	307	1 YC23_SYNY3	P73467 synchocyst
38	46	21.6	346	1 TH13_SCHPO	P36597 schizosacch
39	46	21.6	471	1 SECF_MYCLE	P38386 mycobacteri
40	46	21.6	471	1 YODO_BACSU	O34676 bacillus su
41	46	21.6	692	1 EFG_MYCPU	Q98q98 mycoplasma
42	46	21.6	847	1 SYA_HELPJ	Q9zjy5 helicobacte
43	46	21.6	847	1 SYA_HELPY	P56452 helicobacte
44	46	21.6	923	1 PH87_YEAST	P25360 saccharomyc
45	46	21.6	1036	1 Y414_MYCGE	P47653 mycoplasma

#### ALIGNMENTS

##### RESULT 1

ID	AAKI_HUMAN	STANDARD;	PRT;	464 AA.
AC	Q9UG19; Q9NRL1;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)			
DE	(AMPK gamma3).			
GN	PRKAG3 OR AMPKG3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20164049; PubMed=10698692;			
RA	Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;			
RT	"Characterization of AMP-activated protein kinase gamma-subunit			
RT	isoforms and their role in AMP binding."			
RL	Biochem. J. 346:659-669(2000).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RX	TISSUE=Skeletal muscle;			
RA	MEDLINE=20280150; PubMed=10818001;			
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,			
RA	Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,			
RA	Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,			
RA	Andersson L.;			
RT	"A mutation in PRKAG3 associated with excess glycogen content in pig			
RT	skeletal muscle."			
RL	Science 288:1248-1251(2000).			
CC	-!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID			
CC	SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO			
CC	REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND			
CC	INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-			
CC	SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE			
CC	IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.			
CC	-!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A			
CC	GAMMA NON-CATALYTIC REGULATORY SUBUNITS.			
CC	-!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART			
CC	AND PANCREAS.			
CC	-!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA			
CC	SUBUNIT FAMILY.			
CC	-!- SIMILARITY: CONTAINS 4 CRS DOMAINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AJ249977; CAB65117.1; ALT_INIT.			
DR	EMBL; AF214519; AAF73987.1; -.			

CC modified and this statement is not removed. Usage by and for commerce/  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 CC EMBL; AF214521; AAF73989.1; -.  
 CC EMBL; AF214520; AAF73988.1; -.  
 CC InterPro; IPR000644; CBS\_domain.  
 CC Pfam; PF00571; CBS; 4.  
 CC SMART; SMO0116; CBS; 4.  
 CC Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation.  
 KW DOMAIN 172 226 CBS 1.  
 FT DOMAIN 172 226 CBS 2.  
 FT DOMAIN 253 307 CBS 3.  
 FT DOMAIN 328 381 CBS 3.  
 FT DOMAIN 400 453 CBS 4.  
 FT VARIANT 200 200 R -> Q (IN RN-).  
 FT SEQUENCE 454 AA; 51308 MW; 17639CB12A2BA9DF CRC64;  
 SQ  
 Query Match 98.6%; Score 210; DB 1; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 5.4e-21;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 0;  
 QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||  
 Db 295 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 335  
 -----  
 RESULT 3  
 AAKH HUMAN STANDARD; PRT; 569 AA.  
 ID AAKH HUMAN STANDARD; PRT; 569 AA.  
 AC Q9UGJ0; Q9UND8; Q9NU29; Q9ULX8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-2 subunit (AMPK gamma-2 chain)  
 DE (AMPK gamma2) (H91620p).  
 DE PRKAG2.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A. (ISOFORM A).  
 RX MEDLINE=20164049; PubMed=10698692;  
 RX Cheung P.C.F., Salt I.P., Davies S.P., Hardie D.G., Carling D.;  
 RT "Characterization of AMP-activated protein kinase gamma-subunit  
 RT isoforms and their role in AMP binding.";  
 RT Biochem. J. 346:659-669(2000).  
 RL [2]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RP MEDLINE=20564210; PubMed=11112354;  
 RP Lang T., Yu L., Qiang T., Jiang J., Chen Z., Xin Y., Liu G., Zhao S.;  
 RA "Molecular cloning, genomic organization, and mapping of PRKAG2, a  
 RA heart abundant gamma-2 subunit of 5'-AMP-activated protein kinase, to  
 RA human chromosome 7q36.";  
 RT human genomics 70:258-263(2000).  
 RL [3]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RP TISSUE=Placenta;  
 RC Isoqai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;  
 RT "NEO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RL [4]  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RP TISSUE=Liver;  
 RC Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RL [5]  
 RP SEQUENCE OF 218-569 FROM N.A.  
 RP Hattori A., Seki N., Hayashi A., Kozuma S., Muramatsu M., Saito T.;  
 RA "Human homolog of AMPK gamma-1 chain.";  
 RT







RT Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are  
 RT homologs of proteins that interact with yeast Snf1 protein kinase.";  
 RL J. Biol. Chem. 269:29343-29346(1994).  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HORMONE-SENSITIVE LIPASE AND  
 CC HYDROXYMETHYLGLUTARYL-COA REDUCTASE. APPEARS TO ACT AS A METABOLIC  
 CC STRESS-SENSING PROTEIN KINASE SWITCHING OFF BIOSYNTHETIC PATHWAYS  
 CC WHEN CELLULAR ATP LEVELS ARE DEPLETED AND WHEN 5'-AMP RISES IN  
 CC RESPONSE TO FUEL LIMITATION AND/OR HYPOXIA. THIS IS A REGULATORY  
 CC SUBUNIT.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART AND BRAIN, ALSO  
 CC FOUND IN KIDNEY, WHITE ADIPOSE TISSUE, LUNG AND SPLEEN.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; X95578; CAA64831.1; -;  
 DR EMBL; U42413; AAC52580.1; -;  
 DR InterPro; IPR000644; CBS\_domain.  
 DR Pfam; PF00571; CBS; 4.  
 DR SMART; SM00116; CBS; 4.  
 DR SMART; SM00116; CBS; 4.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT DOMAIN 47 95 CBS 1.  
 FT DOMAIN 122 176 CBS 2.  
 FT DOMAIN 196 249 CBS 3.  
 FT DOMAIN 270 322 CBS 4.  
 FT CONFLICT 114 114 E -> Q (IN REF. 3).  
 FT CONFLICT 201 201 A -> P (IN REF. 3).  
 SQ SEQUENCE 330 AA; 37386 MW; 36031E526C1F1E97 CRC64;  
 Query Match 52.6%; Score 112; DB 1; Length 330;  
 Best Local Similarity 46.3%; Pred. No. 7.3e-08;  
 Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 DB 164 YILTHKRLKFLKFLTFEPKPEFMSKSLLELQIGTYANIA 204  
 [1]  
 RESULT 7  
 AAKG\_HUMAN STANDARD; PRT; 331 AA.  
 AC P54619;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)  
 DE (AMPKg).  
 GN PRKAG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=96224074; PubMed=8621499;  
 RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,  
 RA Dyck J.R.B., Kemp B.E., Witters L.A.;  
 RT "Non-catalytic beta- and gamma-subunit isoforms of the 5'-AMP-activated  
 RT protein kinase.";

RL J. Biol. Chem. 271:8675-8681(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; U42412; AAC50495.1; -;  
 DR EMBL; BC000358; AAH00358.1; -;  
 DR Genew; HGNC:9385; PRKAG1.  
 DR MIM; 602742; -;  
 DR InterPro; IPR000644; CBS\_domain.  
 DR Pfam; PF00571; CBS; 4.  
 DR SMART; SM00116; CBS; 4.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT DOMAIN 48 96 CBS 1.  
 FT DOMAIN 123 177 CBS 2.  
 FT DOMAIN 197 250 CBS 3.  
 FT DOMAIN 271 323 CBS 4.  
 SQ SEQUENCE 331 AA; 37579 MW; 0F22B9CA1DBD87AE CRC64;  
 Query Match 52.6%; Score 112; DB 1; Length 331;  
 Best Local Similarity 46.3%; Pred. No. 7.3e-08;  
 Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 DB 165 YILTHKRLKFLKFLTFEPKPEFMSKSLLELQIGTYANIA 205  
 [1]  
 RESULT 8  
 AAKG\_PIG STANDARD; PRT; 133 AA.  
 AC Q09138;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)  
 DE (AMPKg) (38 kDa subunit) (Fragments).  
 GN PRKAG1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=95050763; PubMed=7961907;  
 RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.,  
 RA Teh T., House C.M., Witters L.A., Kemp B.E.;  
 RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are  
 RT homologs of proteins that interact with yeast Snf1 protein kinase.";  
 RL J. Biol. Chem. 269:29343-29346(1994).  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO

CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.  
 CC -1- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -1- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam: PF00571; CBS; 1.  
 DR SMART: SM00116; CBS; 1.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT NON\_TER 1  
 FT DOMAIN <1 >11 CBS 1.  
 FT DOMAIN <26 >42 CBS 2.  
 FT DOMAIN 62 91 CBS 3.  
 FT DOMAIN <95 127 CBS 4.  
 FT NON\_CONS 11 12  
 FT NON\_CONS 24 25  
 FT NON\_CONS 42 43  
 FT NON\_CONS 73 74  
 FT NON\_CONS 80 81  
 FT NON\_CONS 94 95  
 FT NON\_CONS 103 104  
 FT NON\_TER 133 133  
 SQ SEQUENCE 133 AA; 14763 MW; 84C0C3D41E845CEF CRC64;  
 Query Match 30.3%; Score 64.5; DB 1; Length 133;  
 Best Local Similarity 31.7%; Pred. No. 0.076;  
 Matches 13; Conservative 13; Mismatches 8; Indels 7; Gaps 1;  
 QY 1 HILTHKLLKELHFGSLPRSFYRIQDIGIGTFRDLA 41  
 Db 37 YILTKX-----LFITEPKPEFMSKSLEELQIGTVANIA 70  
 RESULT 9  
 A2AP\_HUMAN STANDARD; PRT; 491 AA.  
 AC P08697;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alpha-2-antiplasmin precursor (Alpha-2-plasmin inhibitor) (Alpha-2-PI)  
 DE (Alpha-2-AP).  
 GN SERPINF2 OR PLI OR AAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88139254; PubMed=2830248;  
 RA Tone M., Kikuno R., Kume-Iwaki A., Hashimoto-Gotoh T.;  
 RT "Structure of human alpha 2-plasmin inhibitor deduced from the cDNA  
 RT sequence.";  
 RL J. Biochem. 102:1033-1041(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88320531; PubMed=3166140;  
 RA Hirosewa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;  
 RT "Organization of the human alpha 2-plasmin inhibitor gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6836-6840(1988).  
 RN [3]  
 RP ERRATUM.  
 RA Hirosewa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:1612-1613(1989).  
 RN [4]  
 RP SEQUENCE OF 4-491 FROM N.A.  
 RX MEDLINE=87109313; PubMed=2433286;  
 RA Holmes W.E., Nelles L., Lijnen H.R., Collen D.;  
 RT "Primary structure of human alpha 2-antiplasmin, a serine protease  
 RT inhibitor (serpin).";

RL J. Biol. Chem. 262:1659-1664(1987).  
 RN [5]  
 RP SEQUENCE OF 218-491 FROM N.A.  
 RX MEDLINE=87137400; PubMed=3818581;  
 RA Sumi Y., Nakamura Y., Aoki N., Sakai M., Muramatsu M.;  
 RT "Structure of the carboxyl-terminal half of human alpha 2-plasmin  
 RT inhibitor deduced from that of cDNA.";  
 RL J. Biochem. 100:1399-1402(1986).  
 RN [6]  
 RP SEQUENCE OF 40-491.  
 RX MEDLINE=87275946; PubMed=2440681;  
 RA Lijnen H.R., Holmes W.E., van Hoef B., Wiman B., Rodriguez H.,  
 RA Collen D.;  
 RT "Amino-acid sequence of human alpha 2-antiplasmin.";  
 RL Eur. J. Biochem. 166:565-574(1987).  
 RN [7]  
 RP SEQUENCE OF 40-43.  
 RX MEDLINE=78023887; PubMed=21075;  
 RA Wiman B., Collen D.;  
 RT "Purification and characterization of human antiplasmin, the  
 RT fast-acting plasmin inhibitor in plasma.";  
 RL Eur. J. Biochem. 78:19-26(1977).  
 RN [8]  
 RP SEQUENCE OF 28-52.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93050153; PubMed=1385210;  
 RA Christensen S., Sotttrup-Jensen L.;  
 RT "Bovine alpha 2-antiplasmin. N-terminal and reactive site sequence.";  
 RL FEBS Lett. 312:100-104(1992).  
 RN [9]  
 RP ACTIVE SITES.  
 RX MEDLINE=88290696; PubMed=2456616;  
 RA Potempa J., Shieh B.-H., Travis J.;  
 RT "Alpha-2-antiplasmin: a serpin with two separate but overlapping  
 RT reactive sites.";  
 RL Science 241:699-700(1988).  
 RN [10]  
 RP SEQUENCE OF 481-491, AND SULFATION.  
 RX MEDLINE=87137577; PubMed=2434496;  
 RA Hortin G., Fok K.F., Toren P.C., Strauss A.W.;  
 RT "Sulfation of a tyrosine residue in the plasmin-binding domain of  
 RT alpha 2-antiplasmin.";  
 RL J. Biol. Chem. 262:3082-3085(1987).  
 RN [11]  
 RP VARIANT OKINAWA.  
 RX MEDLINE=90036902; PubMed=2572590;  
 RA Miura O., Sugahara Y., Aoki N.;  
 RT "Hereditary alpha 2-plasmin inhibitor deficiency caused by a  
 RT transport-deficient mutation (alpha 2-PI-Okinaawa). Deletion of Glu137  
 RT by a trinucleotide deletion blocks intracellular transport.";  
 RL J. Biol. Chem. 264:18213-18219(1989).  
 RN [12]  
 RP VARIANT ALPHA-2-PLASMIN INHIBITOR DEFICIENCY M-411, AND VARIANTS V-27;  
 RP W-33 AND K-434.  
 RX MEDLINE=20051147; PubMed=10583218;  
 RA Lind B., Thorsen S.;  
 RT "A novel missense mutation in the human plasmin inhibitor  
 RT (alpha2-antiplasmin) gene associated with a bleeding tendency.";  
 RL Br. J. Haematol. 107:317-322(1999).  
 CC -1- FUNCTION: THE MAJOR TARGETS OF THIS INHIBITOR ARE PLASMIN AND  
 CC TRYPSIN, BUT IT ALSO INACTIVATES CHYMOTRYPSIN.  
 CC -1- DISEASE: Defects in SERPINF2 are the cause of alpha-2-plasmin  
 CC inhibitor deficiency, a disease resulting in severe hemorrhagic  
 CC diathesis.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RA Hungerer C., Troup B., Jahn D.;  
RT "Cloning and regulation of the Pseudomonas aeruginosa hemf gene  
RT encoding oxygen-dependent coproporphyrinogen III oxidase.";  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN-ATCC 15692 / PA01;  
RC MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
CC -|- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =  
CC protoporphyrinogen-IX + 2 CO(2).  
CC -|- COFACTOR: IRON (BY SIMILARITY).  
CC -|- PATHWAY: porphyrin biosynthesis.  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -|- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X85015; CAA59376.1; -;  
DR EMBL: AE004442; AG03414.1; -;  
DR InterPro: IPR001260; Coprogen\_oxidas.  
DR Pfam: PF01218; Coprogen\_oxidas; 1.  
DR PRINTS: PS01073; COPROGNXDASE.  
DR PROSITE: PS01021; COPROGN\_OXIDASE; 1.  
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.  
SQ SEQUENCE 305 AA; 34806 MW; 240BD06FEA37EA83 CRC64;  
  
Query Match 24.9%; Score 53; DB 1; Length 305;  
Best Local Similarity 36.8%; Pred. No. 6.7;  
Matches 14; Conservative 3; Mismatches 15; Indels 6; Gaps 1;  
  
QY 3 LTRKLLKFLHIFGSLLP-----RPSFLYRTIQDLGI 34  
Db 55 LIKGGVNFHVFQSLPPSASAHRPFLAGRFQALGV 92  
  
RESULT 12  
SZ6L\_HUMAN  
ID SZ6L\_HUMAN STANDARD; PRT; 1024 AA.  
AC Q9BYH1; Q9Y2E1; Q9NU13; Q9NU14; Q9NU15; Q95917; Q9Y3J6;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Seizure 6-like protein precursor.  
GN SZ6L OR KIAA0927.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 4).  
RC TISSUE=Brain;  
RX PubMed=11175339;  
RT "A genomic screen for genes upregulated by demethylation and histone  
RT deacetylase inhibition in human colorectal cancer.";  
RL Nat. Genet. 31:141-149(2002).  
CC -|- FUNCTION: Candidate tumor suppressor gene.  
CC -|- ALTERNATIVE PRODUCTS: 4 isoforms; 1, 2, 3 and 4 (shown here);

RT SEZ6L gene at 22q12.1 in a lung cancer cell line.";  
RL Oncogene 19:6251-6260(2000).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORM 3).  
RC TISSUE=Brain;  
RX MEDLINE=99246063; PubMed=10231032;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Fetal brain;  
RA Collins J.E., Huckle E.J.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
[4]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20057165; PubMed=10591208;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
RA Laird G.D., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Martyn I.D., Mashreghy-Mohammadi M., Matthews L.H., McCann O.T.,  
RA McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,  
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chissole S., Murray J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,  
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,  
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,  
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,  
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,  
RA Seroussi E., Franssion I., Tapia I., Bruder C.E., O'Brien K.P.,  
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,  
RA Tiliahun Y., Wright H.;  
RT "The DNA sequence of human chromosome 22.";  
RL Nature 402:489-495(1999).  
[5]  
RN GENE EXPRESSION REGULATION.  
RX PubMed=11992124;  
RA Suzuki H., Gabrielson E., Chen W., Anbazhagan R., Van Engeland M.,  
RA Weijnenberg M.P., Herman J.G., Baylin S.B.;  
RT "A genomic screen for genes upregulated by demethylation and histone  
RT deacetylase inhibition in human colorectal cancer.";  
RL Nat. Genet. 31:141-149(2002).  
CC -|- FUNCTION: Candidate tumor suppressor gene.  
CC -|- ALTERNATIVE PRODUCTS: 4 isoforms; 1, 2, 3 and 4 (shown here);



```

FT CARBOHYD 296 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 310 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 317 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 28 T -> Q (IN REF. 2).
FT CONFLICT 40 Q -> P (IN REF. 2).
FT CONFLICT 43 Q -> E (IN REF. 2).
SQ SEQUENCE 492 AA; 54710 MW; 0755D6FC9B2DF5D CRC64;

Query Match 24.6%; Score 52.5; DB 1; Length 492;
Best Local Similarity 47.2%; Pred. No. 13;
Matches 17; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

Qy 4 THKRLKFLHI-FGSLPRSPFLYRTIQDLIGTFR 38
Db 129 TLQRLKEVLHADGCPCLPH--LLSLRCQDLGPGAFA 162

RESULT 14
YB64_YEAST
ID YB64_YEAST STANDARD; PRT; 527 AA.
AC P38314;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 57.2 kDa protein in MET8-HEC2 intergenic region.
GN YBR214W OR YBR1501.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Rieger M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 114-527 FROM N.A.
RC STRAIN=S288c;
RA Dubois E., el Bakkoury M., Glandsdorff N., Messenguy F., Pierard A.,
RA Scherens B., Vierendeels F.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YGL056C AND S. POMBE SDS23.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z36083; CAAB5178.1; -.
CC PIR; S46088; S46088.
CC SGD; S0000418; SDS24.
CC InterPro; IPR000644; CBS_domain.
CC Pfam; PF00571; CBS; 4.
CC SMART; SM00116; CBS; 2.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 196 251 CBS 1.
FT DOMAIN 283 335 CBS 2.
SQ SEQUENCE 527 AA; 57187 MW; DC2741550A69C154 CRC64;

Query Match 24.2%; Score 51.5; DB 1; Length 527;
Best Local Similarity 42.1%; Pred. No. 19;
Matches 16; Conservative 6; Mismatches 9; Indels 7; Gaps 2;

Qy 2 ILTHRLKFLF-----HFGSLPRSPFLYRTIQDLIGT 35
Db 240 ILSQRRLIKYLDNARGFTSLEP---LLNSSLQDLHIG 274

RESULT 15

```

```

HEM6_YERPE
ID HEM6_YERPE STANDARD; PRT; 309 AA.
AC Q82CF9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMF OR YP03032.
OS Versinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Farraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -!- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC coproporphyrinogen-IX + 2 CO(2).
CC -!- COFACTOR: Iron (By similarity).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ241455; CAC92274.1; -.
CC InterPro; IPR001260; Coprogen_oxidase.
CC Pfam; PF01218; Coprogen_oxidase; 1.
CC PRINTS; PR00073; COPRGNXDASE.
CC PROSITE; PS01021; COPROGEN_OXIDASE; 1.
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.
SQ SEQUENCE 309 AA; 35004 MW; A3DE259C4C3714D9 CRC64;

Query Match 23.9%; Score 51; DB 1; Length 309;
Best Local Similarity 34.4%; Pred. No. 13;
Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

Qy 9 LKFLHFGSLLP-----RPSFLYRTIQDLGI 34
Db 62 VNFHVSGAMLPASATAHPELAGRSFQALGV 93

Search completed: June 6, 2003, 11:01:31
Job time : 2.50788 secs

```

Result No.	Query	⌈			ID	Description
		Score	Match	Length		
1	124	58.2	566	11	Q91WG5	Q91wg5 mus musculu
2	110	51.6	372	5	Q9BH16	Q9bh16 caenorhabdi
3	105	49.3	634	5	Q9VDD2	Q9vdd2 drosophila
4	105	49.3	647	5	Q9G613	Q9g613 drosophila
5	105	49.3	906	5	Q9SX18	Q9sxt8 drosophila
6	105	49.3	1400	5	Q9SZS7	Q9szs7 drosophila
7	97	45.5	423	5	Q02168	Q02168 caenorhabdi
8	84	39.4	448	5	Q9N501	Q9n501 caenorhabdi
9	66	31.0	1055	12	Q91TR2	Q91tr2 tupaia herp
10	65.5	30.8	577	5	Q8SSV7	Q8ssv7 dictyosteli
11	63	29.6	379	10	Q41108	Q41108 phaseolus v
12	56	26.3	1268	5	Q95NR8	Q95nr8 caenorhabdi
13	56	26.3	1270	5	Q62462	Q62462 caenorhabdi
14	55.5	26.1	516	16	Q9KQK3	Q9kqx3 vibrio chol
15	55	25.8	798	2	Q08329	Q08329 bacillus st
16	54	25.4	71	4	Q9BV31	Q9bv31 homo sapien



```

DE Y11B2A.8 protein.
GN Y11B2A.8
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=990613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; ALI32904; CAC3636.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 372 AA; 41376 MW; 81A39670877167DF CRC64;

Query Match 51.6%; Score 110; DB 5; Length 372;
Best Local Similarity 52.5%; Pred. No. 3.4e-07;
Matches 21; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 40
:|||||:||||: |||||: |::|::|:|:
DB 136 YILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 175

RESULT 3
Q9VDD2 PRELIMINARY; PRT; 634 AA.
ID Q9VDD2
AC Q9VDD2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SNF4AGAMMA protein.
GN SNF4A-GAMMA OR SNF4AGAMMA OR CG17299.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananadides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ihegwam C.,
RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

```

RA Merkurov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003733; AAF5864.1; -.
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 634 AA; 70174 MW; 3FFD0D53E54BBE7C CRC64;

Query Match 49.3%; Score 105; DB 5; Length 634;
Best Local Similarity 40.0%; Pred. No. 2.9e-06;
Matches 16; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 40
:|||||:||||: |||||:|::|::|:|:
DB 293 YILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 332

RESULT 4
Q96613 PRELIMINARY; PRT; 647 AA.
ID Q96613
AC Q96613
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SNF4/AMP-activated protein kinase gamma subunit.
GN SNF4A-GAMMA OR CG17299.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshida E.N., Benkel B.F., Fong Y., Hickey D.A.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF094764; AAC95306.1; -.
DR EMBL; AF094763; AAC95305.1; -.
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Kinase.
SQ SEQUENCE 647 AA; 71592 MW; B792BE1089730B52 CRC64;

Query Match 49.3%; Score 105; DB 5; Length 647;
Best Local Similarity 40.0%; Pred. No. 3e-06;
Matches 16; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

QY 1 HILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 40
:|||||:||||: |||||:|::|::|:|:
DB 306 YILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDL 345

RESULT 5
Q8SXT8 PRELIMINARY; PRT; 906 AA.
ID Q8SXT8
AC Q8SXT8
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

```





```
DR EMBL; AL021483; CAC42378.1; JOINED.
DR EMBL; AL021483; CAC42374.1; -.
DR EMBL; AL021488; CAC42374.1; JOINED.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004182; GRAM_dom.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF02893; GRAM; 2.
DR Pfam; PF00566; TBC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN.1.
SQ SEQUENCE 1268 AA; 144714 MW; 6D9FAA919E19CB16 CRC64;

Query Match 26.3%; Score 56; DB 5; Length 1268;
Best Local Similarity 43.6%; Pred. No. 43;
Matches 17; Conservative 3; Mismatches 11; Indels 8; Gaps 2;

QY 11 FLHIFGSLPRP---SFLYRT-----IQDLGIGTFRDLA 41
   | : | | | | | : | : | : | | | | |
Db 895 FSEVPRLLPWPVTNFIIRVFRLLDISDNGLLTFRDLA 933

RESULT 13
O62462
ID O62462 PRELIMINARY; PRT; 1270 AA.
AC O62462;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Y45F10A.6 protein.
GN Y45F10A.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA McMurray A.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Saldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; AL021488; CAA16368.1; -.
DR EMBL; AL021483; CAA16368.1; JOINED.
DR EMBL; AL021483; CAA16349.1; -.
DR EMBL; AL021488; CAA16349.1; JOINED.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004182; GRAM_dom.
DR Pfam; PF02893; GRAM; 2.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN.1.
SQ SEQUENCE 1270 AA; 144948 MW; ED95EBA90E4D8115 CRC64;

Query Match 26.3%; Score 56; DB 5; Length 1270;
Best Local Similarity 43.6%; Pred. No. 44;
Matches 17; Conservative 3; Mismatches 11; Indels 8; Gaps 2;
```

```
QY 11 FLHIFGSLPRP---SFLYRT-----IQDLGIGTFRDLA 41
   | : | | | | | : | : | : | | | | |
Db 895 FSEVPRLLPWPVTNFIIRVFRLLDISDNGLLTFRDLA 933

RESULT 14
Q9KQX3
ID Q9KQX3 PRELIMINARY; PRT; 516 AA.
AC Q9KQX3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein VC1874.
GN VC1874.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDaniel L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004263; AAF95022.1; -.
DR TIGR; VC1874; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 516 AA; 60882 MW; 18F2C1365D46DFF7 CRC64;

Query Match 26.1%; Score 55.5; DB 16; Length 516;
Best Local Similarity 31.8%; Pred. No. 20;
Matches 14; Conservative 9; Mismatches 16; Indels 5; Gaps 1;

QY 2 ILTHKLLKFLHIFGSLPRPSFLYRTIQD-----LGIGTFRDL 40
   | : | | | | | : | : | : | | | | |
Db 305 LLSDKFMLEFLHSHTSVVAQPPYNSRYFSGINPYALGFAMFRDI 348

RESULT 15
O08329
ID O08329 PRELIMINARY; PRT; 798 AA.
AC O08329;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycogen phosphorylase (EC 2.4.1.1) (Alpha-glucan phosphorylase)
DE (Starch phosphorylase).
GN GLGP.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Geobacillus
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RX MEDLINE=97386405; PubMed=9244254;
RA Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RT "Characterization of a gene cluster for glycogen biosynthesis and a
RT heterotetrameric ADP-glucose pyrophosphorylase from Bacillus
RT stearothermophilus.";
RL J. Bacteriol. 179:4689-4698(1997).
CC -!- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 2.44133 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_320\_360  
Perfect score: 213  
Sequence: 1 HILTHKRLKLFHFGSLP.....PSFLYRTQDLGIGTFRDLA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2.6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	58.2	328	2	US-08-878-989-7
2	124	58.2	328	4	US-09-272-796-7
3	112	52.6	330	4	US-09-359-161-6
4	112	52.6	331	2	US-08-878-989-21
5	112	52.6	331	3	US-09-101-146-64
6	112	52.6	331	4	US-09-272-796-21
7	63	29.6	379	4	US-09-359-161-5
8	54.5	25.6	179	6	5463025-4
9	54.5	25.6	464	6	5463025-1
10	54	25.4	92	4	US-08-905-223-497
11	52.5	24.6	416	4	US-09-330-611-8
12	51	23.9	808	4	US-09-134-001C-3105
13	49	23.0	483	4	US-08-924-183-9
14	49	23.0	483	4	US-09-488-364-9
15	49	23.0	712	1	US-08-587-889-2
16	49	23.0	712	2	US-08-980-060-5
17	49	23.0	712	4	US-09-307-185-5
18	49	23.0	712	5	PCT-US96-09193-2
19	48	22.5	320	4	US-09-576-160B-4
20	48	22.5	350	4	US-09-576-160B-5
21	47.5	22.3	3289	2	US-08-477-451-2
22	47	22.1	314	4	US-09-188-930-193
23	47	22.1	316	4	US-09-188-930-337
24	46.5	21.8	352	4	US-09-576-160B-1
25	46.5	21.8	352	4	US-09-576-160B-2
26	46.5	21.8	352	4	US-09-576-160B-3
27	46.5	21.8	353	1	US-08-229-418-2

28 46.5 21.8 353 2 US-08-932-761A-2  
29 46.5 21.8 353 4 US-09-307-912-2  
30 46.5 21.8 353 5 PCT-US95-04464-2  
31 46 21.6 260 4 US-09-134-001C-4507  
32 46 21.6 471 4 US-09-330-611-10  
33 46 21.6 652 1 US-08-261-663A-6  
34 46 21.6 652 4 US-09-357-206A-5  
35 46 21.6 652 5 PCT-US95-07754A-6  
36 46 21.6 829 1 US-08-346-455B-34  
37 46 21.6 829 3 US-08-977-221-34  
38 46 21.6 829 4 US-09-483-831B-34  
39 46 21.6 829 5 PCT-US95-06613-34  
40 46 21.6 915 1 US-08-346-455B-69  
41 46 21.6 915 3 US-08-977-221-69  
42 46 21.6 915 4 US-09-483-831B-69  
43 46 21.6 915 5 PCT-US95-06613-69  
44 46 21.6 1143 2 US-08-310-912A-108  
45 46 21.6 1143 4 US-09-301-085-108

#### ALIGNMENTS

RESULT 1  
US-08-878-989-7  
; Sequence 7, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PENITUT01







```
; APPLICANT: Dahal, Pectambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; TITLE OF INVENTION: to Stress Conditions in Plants
; FILE REFERENCE: 023070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
; FEATURE:
; OTHER INFORMATION: Phaseolus vulgaris L. Pv42
US-09-359-161-5

Query Match      29.6%; Score 63; DB 4; Length 379;
Best Local Similarity 48.6%; Pred. No. 0.29;
Matches 17; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 2 ILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGT 36
   :|| :||||| 1: | | :||||| |
Db 186 MLTQMDMLKFLHGGGAEL--HSILSRVQDLGADT 218

RESULT 8
5463025-4
; Patent No. 5463025
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 4
; LENGTH: 179
5463025-4

Query Match      25.6%; Score 54.5; DB 6; Length 179;
Best Local Similarity 48.6%; Pred. No. 2.1;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 4 THKRLKFLHIFGSLPRPSFLYRTIQDLGIGT 38
   | :|| : || || | | | | | | | |
Db 129 TLQRLQOVLIHA-GSGPCLPHLLSLRLCQDLGPGAFR 162

RESULT 9
5463025-1
; Patent No. 5463025
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
```

```
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 1
; LENGTH: 464
5463025-1

Query Match      25.6%; Score 54.5; DB 6; Length 464;
Best Local Similarity 48.6%; Pred. No. 6;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 4 THKRLKFLHIFGSLPRPSFLYRTIQDLGIGT 38
   | :|| : || || | | | | | | | |
Db 101 TLQRLQOVLIHA-GSGPCLPHLLSLRLCQDLGPGAFR 134

RESULT 10
US-08-905-223-497
; Sequence 497, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 497:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -26...-1
; IDENTIFICATION METHOD: von Heijne matrix
; OTHER INFORMATION: score 3.6
; OTHER INFORMATION: seq ARSLQFLRLVGQ/LK
US-08-905-223-497

Query Match      25.4%; Score 54; DB 4; Length 92;
Best Local Similarity 40.7%; Pred. No. 1.2;
Matches 11; Conservative 7; Mismatches 7; Indels 2; Gaps 1;

QY 6 KRLKFLHIFGSL--LPRPSFLYRTIQ 30
   : || : || : | : || : || : |
Db 15 RSLQFLRLVGQLKXKVPRTGWIYVNVQ 41
```

```
RESULT 11
US-09-330-611-8
; Sequence 8, Application US/09330611
; Patent No. 6248874
; GENERAL INFORMATION:
; APPLICANT: FREY, Perry A.
; TITLE OF INVENTION: DNA MOLECULES ENCODING BACTERIAL LYSINE 2,3-AMINOMUTASE
; FILE REFERENCE: 032026/0476
; CURRENT APPLICATION NUMBER: US/09/330,611
; CURRENT FILING DATE: 1999-06-11
; EARLIER APPLICATION NUMBER: US 09/198,942
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-330-611-8

Query Match      24.6%; Score 52.5; DB 4; Length 416;
Best Local Similarity 40.5%; Pred. No. 10;
Matches 15; Conservative 5; Mismatches 12; Indels 5; Gaps 2;

QY      4 TH--KRLKLFHFGSLPRPSFLYRTIQDLGIGTGR 38
      || ||| :|: : || :|| ||| |||
Db      270 THVKRL---VHLLVKRVRYIYVCDLSIGIGHGR 303

RESULT 12
US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: CTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3105
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3105

Query Match      23.9%; Score 51; DB 4; Length 808;
Best Local Similarity 30.2%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 11; Indels 10; Gaps 2;

QY      2 ILTHKRLKLFH-----IFGSLPRPSFLYRTIQDLGIGTF 37
      || :|| || || :|| ||| :|| :||
Db      1 ILKEENVSLTHIAVLLPLIFALIP---FLYFVRRIHLGWF 40

RESULT 13
US-08-924-183-9
; Sequence 9, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
```

```
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-08-924-183-9

Query Match      23.0%; Score 49; DB 4; Length 483;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      21 RPSFLYRTIQDLGIGTGRDL 40
      :| ||| :||| ||| :|
Db      19 QPESLYRVVQTLGEGAFGEV 38

RESULT 14
US-09-488-364-9
; Sequence 9, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1013
; CURRENT APPLICATION NUMBER: US/09/488,364
; CURRENT FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-488-364-9

Query Match      23.0%; Score 49; DB 4; Length 483;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      21 RPSFLYRTIQDLGIGTGRDL 40
      :| ||| :||| ||| :|
Db      19 QPESLYRVVQTLGEGAFGEV 38

RESULT 15
US-08-587-889-2
; Sequence 2, Application US/08587889
; Patent No. 5654397
; GENERAL INFORMATION:
; APPLICANT: CAO, Zhaodan
; APPLICANT: CROSTON, Glenn E.
; APPLICANT: GOEDEL, David V.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,889
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

; NAME: Osman Ph.D., Richard Aron  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-60916  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-587-889-2

Query Match 23.0%; Score 49; DB 1; Length 712;  
Best Local Similarity 37.1%; Pred. No. 59;  
Matches 13; Conservative 3; Mismatches 7; Indels 12; Gaps 1;

QY 1 HILTHKRLKPLHIF-----GSLPRPS 23  
Db 85 HILTHQLLRARDIITAWHPPAPLPSPGTAPRPS 119

Search completed: June 6, 2003, 11:04:15  
Job time : 3.44133 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	213	100.0	489	10	US-09-926-581-6		Sequence 6, Appli
2	112	52.6	344	10	US-09-925-297-461		Sequence 461, App
3	105	49.3	1207	9	US-10-108-605-71		Sequence 71, Appl
4	54	25.4	264	9	US-09-510-332-47		Sequence 47, Appl
5	53	24.9	1023	9	US-10-028-072-200		Sequence 200, App
6	53	24.9	1023	9	US-10-121-049-200		Sequence 200, App
7	53	24.9	1023	9	US-10-123-904-200		Sequence 200, App
8	53	24.9	1023	9	US-10-140-470-200		Sequence 200, App
9	53	24.9	1023	9	US-10-175-746-200		Sequence 200, App
10	53	24.9	1023	9	US-10-176-918-200		Sequence 200, App
11	53	24.9	1023	9	US-10-176-921-200		Sequence 200, App
12	53	24.9	1023	9	US-10-137-865-200		Sequence 200, App
13	53	24.9	1023	9	US-10-140-474-200		Sequence 200, App
14	53	24.9	1023	9	US-10-142-431-200		Sequence 200, App
15	53	24.9	1023	9	US-10-143-114-200		Sequence 200, App
16	53	24.9	1023	9	US-10-140-002-200		Sequence 200, App
17	53	24.9	1023	9	US-10-142-419-200		Sequence 200, App
18	53	24.9	1023	9	US-10-123-262-200		Sequence 200, App
19	53	24.9	1023	9	US-10-142-423-200		Sequence 200, App



2	PRIOR FILING DATE: 1998-02-27	
2	PRIOR APPLICATION NUMBER: 60/079728	
2	PRIOR FILING DATE: 1998-03-27	
2	PRIOR APPLICATION NUMBER: 60/080165	
2	PRIOR FILING DATE: 1998-03-31	
2	PRIOR APPLICATION NUMBER: 60/081203	
2	PRIOR FILING DATE: 1998-04-09	
2	PRIOR APPLICATION NUMBER: 60/081229	
2	PRIOR FILING DATE: 1998-04-09	
2	PRIOR APPLICATION NUMBER: 60/081695	
2	PRIOR FILING DATE: 1998-04-14	
2	PRIOR APPLICATION NUMBER: 60/081817	
2	PRIOR FILING DATE: 1998-04-15	
2	PRIOR APPLICATION NUMBER: 60/081818	
2	PRIOR FILING DATE: 1998-04-15	
2	PRIOR APPLICATION NUMBER: 60/082999	
2	PRIOR FILING DATE: 1998-04-24	
2	PRIOR APPLICATION NUMBER: 60/083322	
2	PRIOR FILING DATE: 1998-04-28	
2	PRIOR APPLICATION NUMBER: 60/083545	
2	PRIOR FILING DATE: 1998-04-29	
2	PRIOR APPLICATION NUMBER: 60/084600	
2	PRIOR FILING DATE: 1998-05-07	
2	PRIOR APPLICATION NUMBER: 60/084627	
2	PRIOR FILING DATE: 1998-05-07	
2	PRIOR APPLICATION NUMBER: 60/084637	
2	PRIOR FILING DATE: 1998-05-07	
2	PRIOR APPLICATION NUMBER: 60/085149	
2	PRIOR FILING DATE: 1998-05-12	
2	PRIOR APPLICATION NUMBER: 60/085323	
2	PRIOR FILING DATE: 1998-05-13	
2	PRIOR APPLICATION NUMBER: 60/085338	
2	PRIOR FILING DATE: 1998-05-13	
2	PRIOR APPLICATION NUMBER: 60/085339	
2	PRIOR FILING DATE: 1998-05-13	
2	PRIOR APPLICATION NUMBER: 60/085579	
2	PRIOR FILING DATE: 1998-05-15	
2	PRIOR APPLICATION NUMBER: 60/085697	
2	PRIOR FILING DATE: 1998-05-15	
2	PRIOR APPLICATION NUMBER: 60/085704	
2	PRIOR FILING DATE: 1998-05-15	
2	PRIOR APPLICATION NUMBER: 60/086414	
2	PRIOR FILING DATE: 1998-05-22	
2	PRIOR APPLICATION NUMBER: 60/086430	
2	PRIOR FILING DATE: 1998-05-22	
2	PRIOR APPLICATION NUMBER: 60/087106	
2	PRIOR FILING DATE: 1998-05-28	
2	PRIOR APPLICATION NUMBER: 60/088026	
2	PRIOR FILING DATE: 1998-06-04	
2	PRIOR APPLICATION NUMBER: 60/088730	
2	PRIOR FILING DATE: 1998-06-10	
2	PRIOR APPLICATION NUMBER: 60/088741	
2	PRIOR FILING DATE: 1998-06-10	
2	PRIOR APPLICATION NUMBER: 60/088810	
2	PRIOR FILING DATE: 1998-06-10	
2	PRIOR APPLICATION NUMBER: 60/088858	
2	PRIOR FILING DATE: 19/98-06-11	
2	PRIOR APPLICATION NUMBER: 60/089532	
2	PRIOR FILING DATE: 1998-06-17	
2	PRIOR APPLICATION NUMBER: 60/089599	
2	PRIOR FILING DATE: 1998-06-17	
2	PRIOR APPLICATION NUMBER: 60/089907	
2	PRIOR FILING DATE: 1998-06-18	
2	PRIOR APPLICATION NUMBER: 60/089947	
2	PRIOR FILING DATE: 1998-06-19	
2	PRIOR APPLICATION NUMBER: 60/090349	
2	PRIOR FILING DATE: 1998-06-23	
2	PRIOR APPLICATION NUMBER: 60/090429	
2	PRIOR FILING DATE: 1998-06-24	
2	PRIOR APPLICATION NUMBER: 60/090445	
2	PRIOR FILING DATE: 1998-06-24	
2	PRIOR APPLICATION NUMBER: 60/090538	
2	PRIOR FILING DATE: 1998-06-24	







```

; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKLFHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGQVIRSPNTISVYFRTFDDGLGTFQ 383

RESULT 13
US-10-140-474-200
; Sequence 200, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKLFHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGQVIRSPNTISVYFRTFDDGLGTFQ 383

RESULT 14
US-10-142-431-200
; Sequence 200, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKLFHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGQVIRSPNTISVYFRTFDDGLGTFQ 383

RESULT 15
US-10-143-114-200
; Sequence 200, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKLFHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGQVIRSPNTISVYFRTFDDGLGTFQ 383

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:55:08 ; Search time 78.7881 Seconds  
(without alignments)  
827.023 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2338  
Sequence: 1 MPEGLHARRTPSSSLG.....LSDILQALVLSAGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2338	100.0	489	22	PRKAG3. Homo sapi
2	2398	94.5	464	22	Human AMPK gamma s
3	2085	82.2	514	22	Sus scrofa Prkag3
4	2013	79.3	464	23	Pig AMPK gamma sub
5	2011	79.2	464	22	Pig AMPK gamma sub
6	2011	79.2	464	23	Pig wild-type PRKA
7	2010	79.2	464	23	Pig PRKAG3 polymor
8	2007	79.1	464	23	Pig PRKAG3 polymor
9	2005	79.0	464	23	Pig PRKAG3 polymor
10	1551	61.1	305	22	Human AMPK gamma s

11	1550	61.1	305	22	AAE00329	Human Prkag3 V40I
12	1547	61.0	305	22	AAE00328	Human Prkag3 R41Q
13	1507	59.4	305	22	AAE00220	Pig AMPK gamma sub
14	1506	59.3	305	22	AAE00226	Sus scrofa AMPK ga
15	1503	59.2	305	22	AAE00225	Sus scrofa AMPK ga
16	1080.5	42.6	328	20	AAW88438	Disease associated
17	1080.5	42.6	328	20	AAW93432	Disease associated
18	1017	40.1	331	18	AAW29817	Human protein sequ
19	1017	40.1	331	18	AAW29817	Mammalian AMPK-gam
20	974.5	38.4	353	23	ABW06101	Human pancreatic c
21	886	34.9	488	22	ABG20078	Human NS protein s
22	869.5	34.3	634	22	ABG20078	Novel human diagno
23	845.5	33.3	383	22	ABG20080	Drosophila melanog
24	758	29.9	181	22	ABW11241	Novel human diagno
25	494	19.5	180	22	ABG75798	Human AMP-activate
26	485	19.1	318	21	AAW96788	Human colon cancer
27	396	15.6	149	23	ABW89277	Soybean sucrose no
28	354.5	14.0	492	21	AAW96787	Human polypeptide
29	344.5	13.6	482	21	AAW96786	Soybean sucrose no
30	340	13.4	451	21	AAW96783	Soybean sucrose no
31	326	12.8	493	21	AAW96784	Partial Z. mays su
32	322	12.7	442	21	AAW96789	Rice sucrose non-f
33	316.5	12.5	368	23	ABW90919	Wheat sucrose non-
34	312	12.3	101	21	AAW01655	Herbicideally activ
35	201	7.9	415	21	AAW016812	Human secreted pro
36	201	7.9	424	21	AAW016811	Arabidopsis thalia
37	198	7.8	415	21	AAW54283	Arabidopsis thalia
38	198	7.8	424	21	AAW54282	Arabidopsis thalia
39	193.5	7.6	395	21	AAW16813	Arabidopsis thalia
40	190.5	7.5	122	21	AAW96792	Partial soybean SN
41	190.5	7.5	395	21	AAW54284	Arabidopsis thalia
42	190	7.5	373	22	AAW62807	Tomato LeSNF4 (suc
43	189.5	7.5	352	21	AAW09485	Arabidopsis thalia
44	189.5	7.5	352	21	AAW43443	Arabidopsis thalia
45	189.5	7.5	352	23	ABW91010	Herbicideally activ

ALIGNMENTS

RESULT 1	
AAW47679	AAW47679 standard; Protein; 489 AA.
ID	AAW47679
AC	AAW47679;
XX	
XX	
DT	21-JAN-2002 (first entry)
XX	
DE	PRKAG3.
XX	
KW	Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW	metabolic disease; diabetes; obesity; substitution; ss.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 71
FT	/note= "Possible variation point P71A"
FT	Misc-difference 340
FT	/note= "Possible variation point R340W"
XX	
PN	W0200177305-A2.
XX	
PD	18-OCT-2001.
XX	
PF	06-APR-2001; 2001WO-SE00765.
XX	
PR	07-APR-2000; 2000US-195665P.
XX	
PA	(AREX-) AREXIS AB.
XX	
PI	Andersson L, Luthman H, Marklund S;
XX	



CC of PRKAG3 and is useful in gene therapy.

XX Sequence 464 AA;

SQ Query Match 94.5%; Score 2398; DB 22; Length 464;

Best Local Similarity 99.8%; Pred. No. 2.2e-239;

Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 MSFLEQENSSNPSPAVTSSSIRIRGKRAKALRWTRQKSVEEGEPGQEGPRSPAAE 85

Db 1 MSFLEQENSSNPSPAVTSSSIRIRGKRAKALRWTRQKSVEEGEPGQEGPRSPRPAE 60

QY 86 STGLEATFPKTTPLAQADPAGVPTPTGWDCLPSCDTASAAGSSDDVELATEFPATEAM 145

Db 61 STGLEATFPKTTPLAQADPAGVPTPTGWDCLPSCDTASAAGSSDDVELATEFPATEAM 120

QY 146 ECELEGLLEERFALCLSPQAFPKLGWDDDELKPKGAQIYMRPMQEHYCYDAMATSSKLYI 205

Db 121 ECELEGLLEERFALCLSPQAFPKLGWDDDELKPKGAQIYMRPMQEHYCYDAMATSSKLYI 180

QY 206 FDTMLETKKAFFALVANGVRAAPLWDSKQSFVGMILTITDFILVHLHYRSPVLQIYEIE 265

Db 181 FDTMLETKKAFFALVANGVRAAPLWDSKQSFVGMILTITDFILVHLHYRSPVLQIYEIE 240

QY 266 QHKIETWREIYLOGCFKPLVISPNSLFEAVYTLTKNRIHRLPVLDPVSGNVHLHILTHK 325

Db 241 QHKIETWREIYLOGCFKPLVISPNSLFEAVYTLTKNRIHRLPVLDPVSGNVHLHILTHK 300

QY 326 RLKFLHIFGSLLPSPSFLYRTIODLGIGTFRDLAVVLETPILTALDIFVDRRVSAALPV 385

Db 301 RLKFLHIFGSLLPSPSFLYRTIODLGIGTFRDLAVVLETPILTALDIFVDRRVSAALPV 360

QY 386 VNECGQVGLYSRFDV/IHLAAQQTYNHLDMSVGEALRQRTLCEGLVSCOPHESLGEVID 445

Db 361 VNECGQVGLYSRFDV/IHLAAQQTYNHLDMSVGEALRQRTLCEGLVSCOPHESLGEVID 420

QY 446 RIAREQVHRLVLVDTEQHLLGWSLSDILOALVLSPAGIDALGA 489

Db 421 RIAREQVHRLVLVDTEQHLLGWSLSDILOALVLSPAGIDALGA 464

RESULT 3

AAE00224

XX ID AAE00224 standard; Protein; 514 AA.

XX AC AAE00224;

XX DT 13-JUN-2001 (first entry)

XX DE Sus scrofa Prkag3 splice variant.

XX KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;

KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;

XX cystathione beta synthase; CBS; cardiant; gene therapy.

OS Sus scrofa.

XX WO200120003-A2.

XX PD 22-MAR-2001.

XX PF 11-SEP-2000; 2000WO-EP09896.

XX PR 10-SEP-1999; 99EP-0402236.

PR 18-MAY-2000; 2000EP-0401388.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA (ANDE/) ANDERSSON L.

PA (LOOF/) LOOFT C.

PA (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;

XX WPI; 2001-244810/25.

DR N-PSDB; AAD03321.

XX New variants of the gamma subunit of vertebrate adenosine

PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and

XX myopathy -

PS Claim 5; Page 70-71; 71pp; English.

XX The present sequence is pig adenosine monophosphate (AMP)-activated

CC kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice

CC variant. Prkag3 gene is located in the RN locus of chromosome 15.

CC Mutation in Prkag3 results in an altered regulation of carbohydrate

CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as

CC therapeutic for treating carbohydrate metabolism disorders such as

CC diabetes, obesity, and disorders associated with muscle metabolism

CC such as myopathy and cardiovascular diseases, to modulate AMPK

CC activity, and for restoring a normal AMPK function. PRKAG3 sequence

CC and its functionally altered mutants are useful for the diagnostic

CC evaluation, genetic testing and prognosis of a metabolic disorder,

CC preferably a carbohydrate metabolism disorder. Primers that can detect

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

CC useful for detecting a dysfunction of carbohydrate metabolism resulting

CC from the expression of a functionally altered allele of PRKAG3.

CC Transgenic animal and host cell transformed with PRKAG3 or a

CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

CC screening compounds able to modulate AMPK activity. Nucleic acid

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or

CC in a sequence encoding the first cystathione beta synthase (CBS) domain

CC of PRKAG3 and is useful in gene therapy.

XX SQ Sequence 514 AA;

Query Match 82.2%; Score 2085; DB 22; Length 514;

Best Local Similarity 82.0%; Pred. No. 7.5e-207;

Matches 418; Conservative 20; Mismatches 46; Indels 26; Gaps 3;

QY 5 LEHALRR-----TFSWSSLGSSHEQMSFLEQENSSWPSP 40

Db 6 LEQALRRVPVSGRGWELEQLRPEGRGPTADTFSWSSLGSPKHQMSFLEQESRWPSPR 65

QY 41 AVTSSSERIRGKRAKALRWTRQKSVBEGEPGOGEGPRSPRAAESTGLEATFPKTTPLA 100

Db 66 AVTSSSERHGDGNGKASRWTRQEDVEBEGPPGREGQSPRPVAESTGQEAATFPKATPLA 125

QY 101 QADP-AGVGTPTPTGWDCLPSCDTASAAGSSDDVELATEFPATEAWECELEGLLEERPAL 159

Db 126 QAAPLAEVDPNPPTERDILPSCDCAASASDNTDHLDLGIEFSAASAGDEL-GLVEEKPA 184

QY 160 CLSPQAFPKLGWDDDELKPKGAQIYMRPMQEHYCYDAMATSSKLYIFDTMLEIKKAFAL 219

Db 185 CPSPVELLPRLGWDDDELQPKGAQYVMMHFMQEHYCYDAMATSSKLYIFDTMLEIKKAFAL 244

QY 220 VANGVRAAPLWDSKQSFVGMILTITDFILVHLHYRSPVLQIYEIEQHKIETWREIYLOG 279

Db 245 VANGVRAAPLWDSKQSFVGMILTITDFILVHLHYRSPVLQIYEIEEHKIETWREIYLOG 304

QY 280 CFKPLVSIISPNSLFEAVYTLTKNRIHRLPVLDPVSGNVHLHILTHKRLKFLHIFGSLP 339

Db 305 CFKPLVSIISPNSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLKFLHIFGTLTP 364

QY 340 RPSELYRTIODLGIGTFRDLAVVLETPILTALDIFVDRRVSAALPVVNECQVVGYSRF 399

Db 365 RPSELYRTIODLGIGTFRDLAVVLETPILTALDIFVDRRVSAALPVVNETQVVGYSRF 424

QY 400 DVIHLAAQQTYNHLDMSVGEALRQRTLCEGLVSCOPHESLGEVIDRIARQVHRLVLVD 459

Db 425 DVIHLAAQQTYNHLDMSVGEALRQRTLCEGLVSCOPHETLGEVIDRIARQVHRLVLVD 484

QY 460 ETQHLLGVVSLSDILOALVLSPAGIDALGA 489

```

DB      485 ETQHLLGVVSLDIQLALVLSPPAGIDALGA 514
|||||
RESULT 4
AAE22985
ID      AAE22985 standard; Protein: 464 AA.
XX
AC      AAE22985;
XX
DT      09-AUG-2002 (first entry)
XX
DE      Pig PRKAG3 polymorphic variant (PRKAG3-30).
XX
KW      AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
KW      screening; meat quality; single nucleotide polymorphism; SNP; pig;
KW      variant.
XX
OS      Sus scrofa.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 30
FT      /note= "Wild type Asn is substituted with Thr due
FT      to single nucleotide polymorphism (SNP)"
XX
PN      WO200220850-A2.
XX
PD      14-MAR-2002.
XX
PF      10-SEP-2001; 2001WO-US28283.
XX
PR      08-SEP-2000; 2000US-231045P.
PR      08-JAN-2001; 2001US-260239P.
PR      18-JUN-2001; 2001US-299111P.
XX
PA      (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
PI      Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
DR      WPI; 2002-393850/42.
DR      N-PSDB; AAD36457.
XX
PT      Screening animals to determine those likely to produce larger litters
PT      and improved meat quality traits involves assaying for the presence of
PT      polymorphisms in the AMP activated protein kinase regulatory gamma
PT      subunit gene
XX
PS      Disclosure; Page 91-93; 109pp; English.
XX
CC      The invention relates to a method for screening animals to determine
CC      those more likely to produce large litters and improved meat quality
CC      traits. The method involves assaying for the presence of a genotype
CC      in the sample of genetic material obtained from animal. The genotype
CC      is characterised by polymorphism(s) in the AMP activated protein
CC      kinase regulatory gamma subunit (PRKAG3) gene. The method is used
CC      for screening animals e.g., pigs to determine those most likely to
CC      exhibit improved meat quality traits and to produce larger litters.
CC      The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-30).
XX
SQ      Sequence 464 AA;

Query Match 79.3%; Score 2013; DB 23; Length 464;
Best Local Similarity 86.0%; Pred. No. 1.9e-199;
Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;

QY      26 MSFLEQENSSRPSAVTSSSERIGKRRAKALRWTRQKSVSECEPPQCGEGRSPRAE 85
DB      1 MSFLEQGESRPSRAVTTSSSERSHGQGTAKSRWTRQEDVEEGPGPGREGPGSRPAE 60
QY      86 STGLEATFPKATPLAQADP-AGVGTPTTGMDCILPSPDCTASAAGSSTDDEVLATFPFATEA 144
DB      51 STGQATFPKATPLAQAPLAELVDNPTFERDILPSCAASASDSNTDHLDLGIEFSASAA 120

```

```

QY      145 WECELEGLLEERPALCLSPQAPPFKLGDWDELEKPGCAQIYMRFMQEHCTCYDAMATSSKLV 204
DB      121 SGDEL-GLVEEKPAPCPSPEYLLERLGLGDWDELEKPGCAQIYMRFMQEHCTCYDAMATSSKLV 179
QY      205 IFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTITDFILVHLHRYRSPVLQIYEI 264
DB      180 IFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTITDFILVHLHRYRSPVLQIYEI 239
QY      265 EOHKIETWREIYLOGCFKPLVSPNDLSFEAVYTTILIKNRHRLPVLDPVSGNVLHLTH 324
DB      240 EEHKIETWREIYLOGCFKPLVSPNDLSFEAVYTTILIKNRHRLPVLDPVSGAVLHLTH 299
QY      325 KRLLKFLHIFGSLPRPSFLYRTIQDLGIGCTFRLAVLETAPILTALDIFVDRRVSAIP 384
DB      300 KRLLKFLHIFGSLPRPSFLYRTIQDLGIGCTFRLAVLETAPILTALDIFVDRRVSAIP 359
QY      385 VVNECGQVGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGLVSCQPHESLGEVI 444
DB      360 VVNETGQVGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGLVSCQPHETLGEVI 419
QY      445 DRIAREQVHRLVLDVDETQHLGLGVVSLSDIQLALVLSPPAGIDALGA 489
DB      420 DRIAREQVHRLVLDVDETQHLGLGVVSLSDIQLALVLSPPAGIDALGA 464

RESULT 5
AAE00222
ID      AAE00222 standard; Protein: 464 AA.
XX
AC      AAE00222;
XX
DT      13-JUN-2001 (first entry)
XX
DE      Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.
XX
KW      Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW      PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW      genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW      cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW      chromosome 15.
XX
OS      Sus scrofa.
XX
FH      Key Location/Qualifiers
FT      Domain 172..225
FT      /label= CBS
FT      /note= "Cystathione beta synthase domain"
FT      Misc-difference 200
FT      /note= "RN- mutation site"
FT      Domain 253..307
FT      /label= CBS
FT      /note= "cystathione beta synthase domain"
FT      Domain 329..382
FT      /label= CBS
FT      /note= "Cystathione beta synthase domain"
FT      Domain 400..453
FT      /label= CBS
FT      /note= "cystathione beta synthase domain"
XX
PN      WO200120003-A2.
XX
PD      22-MAR-2001.
XX
PF      11-SEP-2000; 2000WO-EF09896.
XX
PR      10-SEP-1999; 99EP-0402236.
PR      18-MAY-2000; 2000EP-0401388.
XX
PA      (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA      (ANDE/) ANDERSSON L.
PA      (LOOF/) LOOFT C.
PA      (KALM/) KALM E.
XX

```

PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 XX Iannuccelli N, Geillin J, Le Roy P, Chardon P;  
 DR WPI: 2001-244810/25.  
 DR N-PSDB; AAD03319.  
 XX  
 PT New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 PS Claim 5; Fig 3; 7lpp; English.  
 XX  
 CC The present sequence is pig adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
 CC complete PRKAG3. PRKAG3 gene is located in the RN locus of chromosome  
 CC 15. Mutation in Prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 79.2%; Score 2011; DB 22; Length 464;  
 Best Local Similarity 86.0%; Pred. No. 3e-199;  
 Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;  
 QY 26 MSFLEQENSSWSPAVTSSERIGKRKAKALRWTRKSVESGEPGQGGPRSPAAE 85  
 DB 1 MSFLEQGESRWSRAVTTSSERSHGQGNKASRWTRQEDVEEGPPGREGPQSRPAE 60  
 QY 86 STGLEATFPKTPPLAQADP-AGVGHPTPCWDCILSDXTASAAGSSIDDVELATEPATEA 144  
 DB 61 STGQATFPKATPPLAQAAPLAAYVNPPTKRIHLISXKAASASDNTDHLGLIEFSASAA 120  
 QY 145 WECELEGLLEBERPALCISPOAPFTKLGMDUHLRKPGAOIYMRPFMQERTCYDAMATSSKLV 204  
 DB 121 SGDEL-GLVEEKPAQCPSPVLLPLGWDDELQKQAGVYHMPQERTCYDAMATSSKLV 179  
 QY 205 IFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMULTITDFILVLRHYRSPVQIYEI 264  
 DB 180 IFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMULTITDFILVLRHYRSPVQIYEI 239  
 QY 265 EGHKTETWREIYLOGCEKPLVSPNDLSLFEAVYTLKNRHLRPLVLDPSGVNVLHLTH 324  
 DB 240 EGHKTETWREIYLOGCEKPLVSPNDLSLFEAVYALIKNRHLRPLVLDPSGAVLHLTH 299  
 QY 325 KRLKFLHFGSLPLRPSFLYRTIIDLIGTFRDLAVVLETAPILTALDIFDVRVRSALP 384  
 DB 300 KRLKFLHFGTLPLRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFDVRVRSALP 359  
 QY 385 VNEGQGVVGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSQPHESIGEVI 444  
 DB 360 VNETQGVVGLYSRFDVIHLAAQQTYNHLDNMVGEALRQRTLCLEGVLSQPHETLGEVI 419  
 QY 445 DRIARQVHRLVDETHQLLGVVLSLSDILQALVLSPGIDALGA 489  
 DB 420 DRIARQVHRLVDETHQLLGVVLSLSDILQALVLSPGIDALGA 464

RESULT 6  
 AAE22984  
 ID AAE22984 standard; Protein; 464 AA.  
 XX  
 AC AAE22984;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig wild-type PRKAG3 protein.  
 XX  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 30 /note= "Wild type Asn is replaced with Thr during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 52 /note= "Wild type Gly is replaced with Ser during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 199 /note= "Wild type Val is replaced with Ile during  
 FT single nucleotide polymorphism (SNP)"  
 FT Misc-difference 200 /note= "Wild type Arg is replaced with Gln during  
 FT single nucleotide polymorphism (SNP)"  
 XX  
 PN WO200220850-A2.  
 PD 14-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001WO-US28283.  
 XX  
 PR 08-SEP-2000; 2000US-231045P.  
 PR 08-JAN-2001; 2001US-260239P.  
 PR 18-JUN-2001; 2001US-299111P.  
 XX  
 PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 XX  
 PI Rothschild MF, Ciobanu DC, Malek M, Plastow G;  
 DR WPI: 2002-393850/42.  
 DR N-PSDB; AAD36456.  
 XX  
 PT Screening animals to determine those likely to produce larger litters  
 PT and improved meat quality traits involves assaying for the presence of  
 PT polymorphisms in the AMP activated protein kinase regulatory gamma  
 PT subunit gene -  
 XX  
 PS Claim 2; Fig 1; 109pp; English.  
 XX  
 CC The invention relates to a method for screening animals to determine  
 CC those more likely to produce large litters and improved meat quality  
 CC traits. The method involves assaying for the presence of a genotype  
 CC in the sample of genetic material obtained from animal. The genotype  
 CC is characterized by polymorphism(s) in the AMP activated protein  
 CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
 CC for screening animals e.g., pigs to determine those most likely to  
 CC exhibit improved meat quality traits and to produce larger litters.  
 CC The present sequence is pig wild-type PRKAG3 protein.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 79.2%; Score 2011; DB 23; Length 464;  
 Best Local Similarity 86.0%; Pred. No. 3e-199;  
 Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;  
 QY 26 MSFLEQENSSWSPAVTSSERIGKRKAKALRWTRKSVESGEPGQGGPRSPAAE 85  
 DB 1 MSFLEQGESRWSRAVTTSSERSHGQGNKASRWTRQEDVEEGPPGREGPQSRPAE 60





```
XX 14-MAR-2002.
XX 10-SEP-2001; 2001WO-US28283.
XX 08-SEP-2000; 2000US-231045P.
XX 08-JAN-2001; 2001US-260239P.
XX 18-JUN-2001; 2001US-299111P.
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX WPI; 2002-393850/42.
XX N-PSDB; RAD36460.
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene -
XX Disclosure; Page 105-107; 109pp; English.
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).
XX Sequence 464 AA;
XX
XX Query Match 79.1%; Score 2007; DB 23; Length 464;
XX Best Local Similarity 85.8%; Pred. No. 7.8e-199;
XX Matches 399; Conservative 20; Mismatches 44; Indels 2; Gaps 2;
XX
XX QY 26 MSFLEQENSSWSPSPAVTSSSRIRGKRRAKALRWTRQKSVEEGPPQGGGPRSPAAE 85
XX Db 1 MSFLEQGESRSPSRVAVTSSSRSHGDOGKNKASRWTRQEDVEEGPPGREGPQSRPVAE 60
XX
XX QY 86 STGLEATFPKTTPLAQADP-AGVGTPTTGMDCPLSDCTASAAGSDTDDELATFPATEA 144
XX Db 61 STGOEATFPKATPLAQAPLAEDVNDPPTERDILPSCASASDSNTDHLGLIEFSASAA 120
XX
XX QY 145 WECELEGLLEERPALCLSPQAPPFKLGWDDLRKPGAOIYMRFMQEHCTCYDAMATSSKLV 204
XX Db 121 SGDEL-GLVEEKPAFCPSPEVLLPRLGWDDLRKPGAOIYMRFMQEHCTCYDAMATSSKLV 179
XX
XX QY 205 IFDTMLEIKKAFKAFVANGVRAAPLWDSKQSFVGMLTITDFILVHLHRYRSPVLQIYEI 264
XX Db 180 IFDTMLEIKKAFKAFVANGVRAAPLWDSKQSFVGMLTITDFILVHLHRYRSPVLQIYEI 239
XX
XX QY 265 EQHKIETWREIYLCQCFPLVSI SPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHLTH 324
XX Db 240 EEHKIEIETWREIYLCQCFPLVSI SPNDSLFEAVYALIKNRIHRLPVLDPVSGVHLHLTH 299
XX
XX QY 325 KRLKLFHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLETAIPILTALDIFVDRVSALP 384
XX Db 300 KRLKLFHIFGTLPRPSFLYRTIQDLGIGTFRDLAVLETAIPILTALDIFVDRVSALP 359
XX
XX QY 395 VVNECGQVGLYSREVDVHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVI 444
XX Db 360 VVNETGVGLYSREVDVHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSQCPHETLGEVI 419
XX
XX QY 445 DRIAREQVHRLVLDVETQHLGVVLSLDILQALVLSPAGIDALGA 489
XX Db 420 DRIAREQVHRLVLDVETQHLGVVLSLDILQALVLSPAGIDALGA 464
XX
XX RESULT 9
```

```
AAE22986
ID AAE22986 standard; Protein; 464 AA.
XX
XX AC AAE22986;
XX
XX DT 09-AUG-2002 (first entry)
XX
XX DE Pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
XX screening; meat quality; single nucleotide polymorphism; SNP; pig;
XX variant.
XX
XX OS Sus scrofa.
XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 52
XX FT /note= "Wild type Gly is substituted with Ser due
XX FT to single nucleotide polymorphism (SNP)"
XX
XX PN WO200220850-A2.
XX
XX PD 14-MAR-2002.
XX
XX PF 10-SEP-2001; 2001WO-US28283.
XX
XX PR 08-SEP-2000; 2000US-231045P.
XX PR 08-JAN-2001; 2001US-260239P.
XX PR 18-JUN-2001; 2001US-299111P.
XX
XX PA (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX PI Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
XX DR WPI; 2002-393850/42.
XX DR N-PSDB; AAD36458.
XX
XX PT Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene -
XX
XX PS Claim 36; Page 96-97; 109pp; English.
XX
XX CC The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX SQ Sequence 464 AA;
XX
XX Query Match 79.0%; Score 2005; DB 23; Length 464;
XX Best Local Similarity 85.8%; Pred. No. 1.3e-198;
XX Matches 399; Conservative 19; Mismatches 45; Indels 2; Gaps 2;
XX
XX QY 26 MSFLEQENSSWSPSPAVTSSSRIRGKRRAKALRWTRQKSVEEGPPQGGGPRSPAAE 85
XX Db 1 MSFLEQGESRSPSRVAVTSSSRSHGDOGKNKASRWTRQEDVEEGPPGREGPQSRPVAE 60
XX
XX QY 86 STGLEATFPKTTPLAQADP-AGVGTPTTGMDCPLSDCTASAAGSDTDDELATFPATEA 144
XX Db 61 STGOEATFPKATPLAQAPLAEDVNDPPTERDILPSCASASDSNTDHLGLIEFSASAA 120
XX
XX QY 145 WECELEGLLEERPALCLSPQAPPFKLGWDDLRKPGAOIYMRFMQEHCTCYDAMATSSKLV 204
XX Db 121 SGDEL-GLVEEKPAFCPSPEVLLPRLGWDDLRKPGAOIYMRFMQEHCTCYDAMATSSKLV 179
XX
XX QY 205 IFDTMLEIKKAFKAFVANGVRAAPLWDSKQSFVGMLTITDFILVHLHRYRSPVLQIYEI 264
XX
XX SQ Sequence 464 AA;
```

Db 180 IFDTMLETKAFFALVANGVRAAPLWDSKKQSFVGMLTITDFILVLRHYRSPVQIYEI 239  
QY 265 EQHKETWREIYLOGCEKPLVSI SPNDSLFEAVYTLIKNRTHRLPVLDPVSGNVHLHILTH 324  
Db 240 EHKETWREIYLOGCEKPLVSI SPNDSLFEAVYALIKNRHLRLPVLDPVSGNAVHLHILTH 299  
QY 325 KRLKFLHIFGSLLPSPSELYRTIQDLGIGTFRDLAVVLETPAIIALTALDIFVDRRVSAALP 384  
Db 300 KRLKFLHIFGSLLPSPSELYRTIQDLGIGTFRDLAVVLETPAIIALTALDIFVDRRVSAALP 359  
QY 385 VVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHESLGEVI 444  
Db 360 VVNETGVVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQRTLCLEGVLSQCPHETILGEVI 419  
QY 445 DRIARQVHRLVYDVTQHLLGVVSLSDILQALVLSAGIDALGA 489  
Db 420 DRIARQVHRLVYDVTQHLLGVVSLSDILQALVLSAGIDALGA 464  
RESULT 10  
AAE00221  
ID AAE00221 standard; Protein; 305 AA.  
XX  
AC AAE00221;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Human AMPK gamma subunit muscle-specific isoform, PRKAG3.  
XX  
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200120003-A2.  
XX  
PD 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-EP09896.  
XX  
PR 10-SEP-1999; 99EP-0402236.  
XX  
PR 18-MAY-2000; 2000EP-0401388.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX  
XX WPI; 2001-244810/25.  
DR N-PSDB; AAD03296.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX  
PS Claim 4; Page 55-57; 71pp; English.  
XX  
XX The present sequence is human adenosine monophosphate  
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
CC PRKAG3. Mutation in Prkag3 results in an altered regulation of  
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
CC useful as therapeutic for treating carbohydrate metabolism disorders such  
CC as diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic

CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
XX  
XX Sequence 305 AA;  
SQ  
Query Match 61.1%; Score 1551; DB 22; Length 305;  
Best Local Similarity 100.0%; Pred. No. 7.7e-152;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 185 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFKAFVANGVRAAPLWDSKKQSFVGMLTIT 244  
Db 1 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFKAFVANGVRAAPLWDSKKQSFVGMLTIT 60  
QY 245 DFILVLRHYRSPVQIYEI EIQHKIETWREIYLOGCEKPLVSI SPNDSLFEAVYTLIKNR 304  
Db 61 DFILVLRHYRSPVQIYEI EIQHKIETWREIYLOGCEKPLVSI SPNDSLFEAVYTLIKNR 120  
QY 305 IHRLPVLDPVSGNVHLHILTHKRLKFLHIFGSLLPSPSELYRTIQDLGIGTFRDLAVVLE 364  
Db 121 IHRLPVLDPVSGNVHLHILTHKRLKFLHIFGSLLPSPSELYRTIQDLGIGTFRDLAVVLE 180  
QY 365 TAPILTALDIFVDRRVSAALPVVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALROR 424  
Db 181 TAPILTALDIFVDRRVSAALPVVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALROR 240  
QY 425 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVYDVTQHLLGVVSLSDILQALVLSAGI 484  
Db 241 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVYDVTQHLLGVVSLSDILQALVLSAGI 300  
QY 485 DALGA 489  
Db 301 DALGA 305  
RESULT 11  
AAE00329  
ID AAE00329 standard; Protein; 305 AA.  
XX  
AC AAE00329;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Human Prkag3 V40I mutant.  
XX  
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 40  
FT /note= "Wild-type Val substituted with Ile"  
XX  
XX WO200120003-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 11-SEP-2000; 2000WO-EP09896.  
XX  
XX 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX  
DR WPI; 2001-244810/25.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX  
PS Cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
PS Disclosure; Page -: 71pp; English.  
XX  
CC The present sequence is a V40I mutant of human muscle-specific isoform  
CC of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
CC (AMPK) Prkag3. This mutant sequence results in decreased glycogen  
CC content in human skeletal muscle.  
CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
CC therapeutic for treating carbohydrate metabolism disorders such as  
CC diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
CC Note: The present sequence is not shown in the specification, but  
CC is derived from the human Prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
CC of sequence listing (AAE00221).  
XX  
SQ Sequence 305 AA:  
Query Match 61.1%; Score 1550; DB 22; Length 305;  
Best Local Similarity 99.7%; Pred. No. 9.8e-152;  
Matches 304; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 185 MRFMQEHTCYDAMATSSKLVFDFTMLEIKKAFALVANGVRAAPLWDSKQSFVGMLTIT 244  
Db 1 MRFMQEHTCYDAMATSSKLVFDFTMLEIKKAFALVANGVRAAPLWDSKQSFVGMLTIT 60  
QY 245 DFILVILHRYRSPVQIYEIQHETWREIYLQGCFFPLVSPNSLSFEAVYTLIKNR 304  
Db 61 DFILVILHRYRSPVQIYEIQHETWREIYLQGCFFPLVSPNSLSFEAVYTLIKNR 120  
QY 305 IHLRLPVLDPVSGNVLHILTHKRLKLFHIFGSLIRPPSFLYRTQDILGIGTFRDLAVLLE 364  
Db 121 IHLRLPVLDPVSGNVLHILTHKRLKLFHIFGSLIRPPSFLYRTQDILGIGTFRDLAVLLE 180  
QY 365 TAPILTALDIFDVRVSALPVNCGVGLYSRFDVTHLAQQTYNHLDMSVGEALROR 424  
Db 181 TAPILTALDIFDVRVSALPVNCGVGLYSRFDVTHLAQQTYNHLDMSVGEALROR 240  
QY 425 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGVSLSLDILQALVSPAGI 484  
Db 241 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGVSLSLDILQALVSPAGI 300  
QY 485 DALGA 489  
Db 301 DALGA 305

RESULT 12  
AAE00328  
ID AAE00328 standard; Protein; 305 AA.  
XX  
AC AAE00328;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Human Prkag3 R41Q mutant.  
XX  
KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW Cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 41  
FT /note= "Wild-type Arg substituted with Gln"  
XX  
PN WO200120003-A2.  
XX  
XX 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-EP09896.  
XX  
PR 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.  
XX  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX  
DR WPI; 2001-244810/25.  
XX  
XX New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX  
PS Disclosure; Page -: 71pp; English.  
XX  
CC The present sequence is a R41Q mutant of human muscle-specific isoform  
CC of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
CC (AMPK) Prkag3. This mutant sequence results in increased glycogen  
CC content in human skeletal muscle.  
CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
CC therapeutic for treating carbohydrate metabolism disorders such as  
CC diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
CC Note: The present sequence is not shown in the specification, but  
CC is derived from the human Prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
CC of sequence listing (AAE00221).  
XX

```
CC is derived from the human Prkg3 sequence SEQ.ID.NO.4 shown in page 57-58
CC of sequence listing (AAE00221).
XX
SQ Sequence 305 AA;

Query Match 61.08; Score 1547; DB 22; Length 305;
Best Local Similarity 99.7%; Pred. No. 2e-151;
Matches 304; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 185 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTIT 244
Db 1 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTIT 60

Qy 245 DFILVLRHYRSPPLVQIYEIEQHKIETWREIYLOGCFKPLVSPNDSLFEAVYTLIKNR 304
Db 61 DFILVLRHYRSPPLVQIYEIEQHKIETWREIYLOGCFKPLVSPNDSLFEAVYTLIKNR 120

Qy 305 IHRPLVLPDPSGVNVLHILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLLE 364
Db 121 IHRPLVLPDPSGVNVLHILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLLE 180

Qy 365 TAPILTALDIFVDRRVSAIPVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 424
Db 181 TAPILTALDIFVDRRVSAIPVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 240

Qy 425 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGLGVVSLSDILQALVLSPAGI 484
Db 241 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGLGVVSLSDILQALVLSPAGI 300

Qy 485 DALGA 489
Db 301 DALGA 305

RESULT 13
AAE00220
ID AAE00220 standard; Protein; 305 AA.
XX
AC AAE00220;
XX
DT 13-JUN-2001 (first entry)
XX
DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3.
XX
KW Pig: gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW chromosome 15.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Domain 13..66
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 94..148
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
```

```
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI: 2001-244810/25.
DR N-PSDB: AAD03295.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy .
XX
PS Claim 4; Fig 2; 71pp; English.
XX
CC The present amino acid sequence is pig adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC PRKAG3. Prkg3 gene is located in the RN locus of chromosome 15.
CC Mutation in Prkg3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkg3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 305 AA;

Query Match 59.4%; Score 1507; DB 22; Length 305;
Best Local Similarity 97.0%; Pred. No. 2.8e-147;
Matches 296; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 185 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTIT 244
Db 1 MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTIT 60

Qy 245 DFILVLRHYRSPPLVQIYEIEQHKIETWREIYLOGCFKPLVSPNDSLFEAVYTLIKNR 304
Db 61 DFILVLRHYRSPPLVQIYEIEQHKIETWREIYLOGCFKPLVSPNDSLFEAVYTLIKNR 120

Qy 305 IHRPLVLPDPSGVNVLHILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLLE 364
Db 121 IHRPLVLPDPSGVNVLHILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVLLE 180

Qy 365 TAPILTALDIFVDRRVSAIPVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 424
Db 181 TAPILTALDIFVDRRVSAIPVNECGVGLYSRFDVIHLAAQOTYNHLDMSVGEALRQR 240

Qy 425 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGLGVVSLSDILQALVLSPAGI 484
Db 241 TLCLEGVLSQCPHESLGEVIDRIAREQVHRLVLVDVETQHLGLGVVSLSDILQALVLSPAGI 300

Qy 485 DALGA 489
Db 301 DALGA 305
```

```

RESULT 14
ID AAE00226 standard; Protein; 305 AA.
XX
AC AAE00226;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).
XX
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;
KW variant.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Domain 13..66
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Misc-difference 40
FT /note= "Wild-type Val substituted with Ile"
FT Domain 94..148
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
WPT; 2001-244810/25.
XX
XX
XX New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 9; Page -; 71pp; English.
XX
XX
XX The present sequence is a V40I mutant of muscle-specific isoform of
CC gamma subunit of adenosine monophosphate (AMP)-activated kinase
CC (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in
CC decreased glycogen content in pig skeletal muscle.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
```

```

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
CC Note: The present sequence is not shown in the specification, but
CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).
XX
SQ Sequence 305 AA;
Query Match 59.3%; Score 1506; DB 22; Length 305;
Best Local Similarity 96.7%; Pred. No. 3.6e-147;
Matches 295; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 185 MRFQEHCTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 244
DB 1 MHFMQEHCTCYDAMATSSKLVIFDTMLEIKKAPFALVANGVRAAPLWDSKKQSFVGMLTIT 60
QY 245 DFILVLRHYRSPVQIYEIEQHKTETWREIYLOGCFPLYSISPNDLSLFEAVYTLIKNR 304
DB 61 DFILVLRHYRSPVQIYEIEEHKTETWREIYLOGCFPLYSISPNDLSLFEAVYALIKNR 120
QY 305 IHRLPVLDPVSGNVHLHILTHKRLKFLHIFGSLPRPSFLYRTIODLGIGTFRDLAVLLE 364
DB 121 IHRLPVLDPVSGAVLHILTHKRLKFLHIFGTLPRPSFLYRTIQLDGLGTFRDLAVLLE 180
QY 365 TAPILTALDIFVDRVRSALPVNCGQVGLYSRFDVTHLAAQOTYNHLDMSVGEALROR 424
DB 181 TAPILTALDIFVDRVRSALPVNCGVGLYSRFDVTHLAAQOTYNHLDMSVGEALROR 240
QY 425 TLCLEGVLSQPHESIGVIDRIAREQVHRLVLVDVDETHLGVLSLSDILQALVLSPAGI 484
DB 241 TLCLEGVLSQPHETLGEVIDRIAREQVHRLVLVDVDETHLGVLSLSDILQALVLSPAGI 300
QY 485 DALGA 489
DB 301 DALGA 305
RESULT 15
AAE00225
ID AAE00225 standard; Protein; 305 AA.
XX
AC AAE00225;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).
XX
XX
XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
XX PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
XX genetic testing; carbohydrate metabolism disorder; skeletal muscle;
XX cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;
XX variant.
XX
OS Sus scrofa.
XX
XX
XX Key Location/Qualifiers
XX FH Domain 13..66
XX /label= CBS
XX /note= "Cystathione beta synthase domain"
XX FT Misc-difference 41
XX /note= "Wild-type Arg substituted with Gln"
XX FT Domain 94..148
XX /label= CBS
XX /note= "Cystathione beta synthase domain"
XX FT Domain 170..223
XX /label= CBS
XX /note= "Cystathione beta synthase domain"
```

```

FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
PN WO200120003-A2.
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PJ Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI; 2001-244810/25.
XX
XX New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy .
XX
PS Claim 9; Page 7; 7lpp; English.
XX
CC The present sequence is a R41Q mutant of muscle-specific isoform of
CC gamma subunit of adenosine monophosphate (AMP)-activated kinase
CC (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in
CC increased glycogen content in pig skeletal muscle.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
CC Note: The present sequence is not shown in the specification, but
CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).
XX
SQ Sequence 305 AA:

Query Match 59.2%; Score 1503; DB 22; Length 305;
Best Local Similarity 96.7%; Pred. No. 7.3e-147;
Matches 295; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 185 MRFMQHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTIT 244
DB 1 MRFMQHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVQAAPLWDSKKQSFVGMLTIT 60
QY 245 DFILVLRHYRSPLOQIYEIOHKIETWREIYLOQCFPLVSI PNDSLFEAVYTLIKNR 304
DB 61 DFILVLRHYRSPLOQIYEIEEHKIETWREIYLOQCFPLVSI PNDSLFEAVYALIKNR 120
QY 305 IHRLPVLDPSVGNVLHILTHKRLKFLHIFGSLIPRPSFLYRTIQDLGIGTFRDLAVLLE 364
DB 121 IHRLPVLDPSGAVLHILTHKRLKFLHIFGTLIPRPSFLYRTIQDLGIGTFRDLAVLLE 180
```

```

QY 365 TABILTALDIFVDRRVVSALPVVNECGQVGLYSRFDVTHLAAQQTYNHLDMSVGEALROR 424
DB TABILTALDIFVDRRVVSALPVVNETGQVWGLYSRFDVTHLAAQQTYNHLDMNYGEALROR 240
QY 425 TLCLEGVLSQPHESLGEVIDRIAREQVHRLVLDVETQHLLGVVSLSDIILQALVLSPAGI 484
DB TLCLEGVLSQPHETLGEVIDRIVREQVHRLVLDVETQHLLGVVSLSDIILQALVLSPAGI 300
QY 485 DALGA 489
DB 301 DALGA 305
```

Search completed: June 6, 2003, 11:01:00  
Job time : 80.7881 secs



T25899  
hypothetical protein T20F7.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T25899  
R:Miller, N.; Gattung, S.  
submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of *C. elegans* cosmid T20F7.  
A:Reference number: T20107  
A:Accession: T25899  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-478 <ML>  
A:Cross-references: EMBL:U97550; PIDN:AAB52856.1; GSPDB:GN00028; CESP:T20F7.6  
A:Experimental source: strain Bristol N2; clone T20F7  
C:Genetics:  
A:Gene: CESP:T20F7.6  
A:Map position: X  
A:Introns: 33/3; 112/2; 144/1; 205/3; 263/2; 300/3; 380/3; 402/2; 451/1

Query Match 19.2%; Score 488; DB 2; Length 478;  
Best Local Similarity 36.7%; Pred. No. 5.4e-30;  
Matches 119; Conservative 69; Mismatches 106; Indels 30; Gaps 8;

QY 182 QIYVRFQCHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSK-KQSFYGM 240  
DB 38 EFAPRLWLNOCYDAMFSSKMWVFDGGLMKAFLGGLAOSTRHVLLSDPFGGKLGGI 97  
QY 241 LATDTFILVLRHYRS-----PIVQI--YEIQKHKEIWRB-IYLOGCKPKPLVS 286  
DB 98 LSWDTDFIKVLMKIVRETKCEKSTELDTQIANEEIGNLSIRQRELYKKEGNLRPLVS 157  
QY 287 ISPNDLSPEANYTLIKRIIRLPLVDPVSGNVLHLTHKRLKPLHIFGSLPRPSFLYR 346  
DB 158 VDASGSLDAACILAEHRVRIHVDPDLSALFTLTHKRLKPLWFGKHLAPLEYLHK 217  
QY 347 TIQDLGIGTFRLAWLETAPILTALDIFVDRRVSAIPV-NEOGVVGVLSRFVDVHLA 405  
DB 218 SPKEIGCWSGIRVPFDQLVCLDLILKNGVSLGVPRFTRFVVDVMSYRFAVGIA 277  
QY 406 AQQTYNHLDMSVGALKRQRTLCLEG-----VLSCQPHESLGEVDIARIAREQVRLVL 457  
DB 278 LE---NRLLDITVKEALAFKS---QGGPMKNDERVSVSRDNEFWKAVNVLVDHNVRLCA 331  
QY 458 VBETHLLGVVSVLSDILOALVLSP 481  
DB 332 VNEHGGIEGVTSLSDVINPMVVQP 355

RESULT 3  
RGBYC3  
regulatory protein SNF4 - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: CAT3 protein; protein G2945; protein YGL115W  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 21-Jul-2000  
C:Accession: A38906; J0316; S48508; S64125  
R:Celenza, J.L.; Eng, F.J.; Carlson, M.  
Mol. Cell. Biol. 9, 5045-5054, 1989  
A:Title: Molecular analysis of the SNF4 gene of *Saccharomyces cerevisiae*: evidence for  
A:Reference number: A33480; MUID:90097921; PMID:2481228  
A:Accession: A38906  
A:Molecule type: DNA  
A:Residues: 1-322 <CEL>  
A:Cross-references: GB:M30470; NID:g172635; PIDN:AAA35061.1; PID:g172636  
R:Schueller, H.J.; Entian, K.D.  
Gene 67, 247-257, 1988  
A:Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucose  
A:Reference number: J0316; MUID:89006284; PMID:3049255  
A:Accession: J0316  
A:Molecule type: DNA  
A:Residues: 1-322 <SCH>  
A:Cross-references: GB:M21760; NID:g171164; PIDN:AAA34472.1; PID:g171165  
R:Doi, A.; Doi, K.

submitted to the EMBL Data Library, June 1993  
A:Description: Correct end of the ORF for the CDC20 gene of *Saccharomyces cerevisiae*  
A:Reference number: S48507  
A:Accession: S48508  
A:Molecule type: DNA  
A:Residues: 1-21 <DOI>  
A:Cross-references: EMBL:D16506; NID:g391938; PIDN:BAA03958.1; PID:g2160324  
R:Lauguin, G.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64122  
A:Accession: S64125  
A:Molecule type: DNA  
A:Residues: 1-322 <LAU>  
A:Cross-references: EMBL:J72637; NID:g1322666; PIDN:CAA96823.1; PID:g1322667; GSPDB:  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:SNF4; CAT34; MIPS:YGL115W  
A:Cross-references: SGD:S0003083; MIPS:YGL115W  
A:Map position: 7L  
C:Function:  
A:Description: involved in derepression of glucose-repressed genes  
C:Superfamily: CAT3 protein  
C:Keywords: nucleus; transcription regulation

Query Match 18.0%; Score 457; DB 1; Length 322;  
Best Local Similarity 34.2%; Pred. No. 7.9e-28;  
Matches 106; Conservative 66; Mismatches 108; Indels 30; Gaps 6;

QY 186 RFMOHFTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKQSFYGMITITD 245  
DB 24 KFLNKSITVDLPVSYRLIVLDTLSLVKKSLLVNLONSIVSAPLWDSKTRFAGLLTTTD 83  
QY 246 FILVLRHYRSPVLQVIEIQKHKEIWRREIYLOGCFKPL-----VTSIPNDSLF 294  
DB 84 FINVIQYYSNP-----DKFELVDKLQLDG-LKDIERALGVQDLDTASIHPSRPLF 133  
QY 295 EAYVTILIKRIIRLPLVDPVSGN---VLHILTHKRLKPLHIFGSLPRPS-FLYRTIQ 349  
DB 134 EACLKMSKGRIPLDIQDEETHREIVSVLTQYRIKLFV----ALNCRETHFLKIPIG 189  
QY 350 DLGIGTFRLAVVLETAPILTALDIFVDRRVSAIPVNECGVVGVLSRFVDVHLAAQQT 409  
DB 190 DLNITQDNKSCQMTTPVDVIGQLTQGVSSVPIIDENGVLINVEAYDVLGLIKGGI 249  
QY 410 YNHLDMSVGALKRQRTLCLEGVLSCQPHESLGEVDIARIAREQVRLVLVDTOHLLGVWS 469  
DB 250 YNDLSLSVGEALMRSDDFCGVYCTCKNDKLSITINDINRKARVHRFVVDVGVGLVGLT 309  
QY 470 LSDLQALVYL 479  
DB 310 LSDLKAYILL 319

RESULT 4  
S67444  
probable 5'-AMP-activated, gamma subunit family - fission yeast (*Schizosaccharomyces*  
C:Species: *Schizosaccharomyces pombe*  
C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000  
C:Accession: T38059; S67444  
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: Z21766  
A:Accession: T38059  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <MC2>  
A:Cross-references: EMBL:Z69944; NID:g1217974; PIDN:CAA93805.1; PID:g1217975; GSPDB:  
A:Introns: 12/3; 54/1; 173/3; 226/1  
C:Superfamily: CAT3 protein

Query Match 14.8%; Score 375.5; DB 2; Length 274;





```
Matches 105; Conservative 87; Mismatches 169; Indels 121; Gaps 20;
QY 62 RQKSVBEGPPQGGCPRS-----RPAAEISGLEATPKTTPLAQADPAGVGTPT 112
Db 3 RLTMLEKGDPTLAGSPKNYSKPKPDTPTASEQRNSTSGRRPSFLSQV--RDLIHPRT 60
QY 113 GWDCPLSDCTASAAGSS-----TDDVELATEFFATEAWCELEGLLEPAPALCISPCA 165
Db 61 G-----SLGASFRPEATRQVQA-----ECNVEAYDLR-----DKHA 95
QY 166 PFPKLGWDE-----LKKPQAQI-----YMRFMOEHCTCYDA 196
Db 96 HF-----YDDNGFRPRNSGDLILIRKTSRPISENVVGVIDQKTDPHYQYMSVVDVYEL 151
QY 197 MATSSKLVLFDPMLEIKKAPALVANGVRAAPLMDSKKQSEV--GMLTITDFTLVLRHY 254
Db 152 CPNNSKVIIDASTPTTFAFRMDHNITTLIVMOTSDARHVKNKNIUTLFDCLNARN-- 209
QY 255 RSLPVOIYEIQHKIETWREIYLGQCFKPLVSI SPNDSIFEAVYTLTKNRIHRLPVDPV 314
Db 210 ETTPADGQVLRASDI-----LSG--NQLSVSISSKILDOLCEELHONLHRVYVLDGA 260
QY 315 SGNVHLHLTKRLKFLHIFOSLLPSPSFHYKTQIDLGIGTFROLAVVLETPALITALDI 374
Db 261 K-EVNNIISVRVIAAIHKQKRSIHFAQWLSKSIQMSAIGTWENVAIVSQNETVYFRMBD 319
QY 375 FVDRRYSALPVNCCGQVGVLYSRFDVHLLAAQOTYNNH-----DMSVGEALRQRTLC 427
Db 320 MLGFHSYALPVVDSQNVGIVYTKTDICKALPR--NFIEPKRWLQTKYSDILH---IC 373
QY 428 LEGVLSQPHESLGEVIDRIAREQVHRLVLVDETO-----HLLGWLSLSDILQALVL 479
Db 374 KSQIL-ISSADSQGVQVLDL-----LAGDTQSAFAIHNGKAIGVISTLDFLSHLR 423
QY 480 SP 481
Db 424 SP 425

RESULT 8
T10971
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: T10971
K:Abc. H.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z17236
A:Accession: T10971
A>Status: preliminary; translated from GH/EMBL/UDBAJ
A:Molecule type: mRNA
A:Residues: 1-379 <ABE>
A:Cross-references: EMBL:U40713; NID:g1113940; PID:g1113941

Query Match 8.4%; Score 214; DB 2; Length 379;
Best Local Similarity 23.5%; Pred. No. 7.le-09;
Matches 85; Conservative 73; Mismatches 122; Indels 82; Gaps 17;
QY 188 MOEHCTYDAMATSSKLVIPTDMLFIKKAFFALVANGVRAAPL-----W-----DS 232
Db 17 LKEKCYKDMVMGKKELEVPVPTASLAQIMNTLVANKIVAVPAAPPQGWAGCGSMIVES 76
QY 233 KKQS-----FVGMELTITDFTLVL-----HRYVRSPLVQIYEIOHKIETWREIY---LQG 279
Db 77 DKQTGAVRKHYYIGWVMTLDILAHITAGDDHLSCGDNITQ-DLDQMSDSVSSIIHGSFEG 134
QY 280 CFKPLVSI SPNDSLFEAVYTLTKNRIHRLPVDLPVSG---NV-----LHLTH 324
Db 135 L--SLWTPNTPNTSMIDCMEVFSKG-VHRAMV--PVDGLEENYASGVSELTESASSYQMLTQ 189
QY 325 KRLKLFHIFGSLPPSPFLYRTQDGLIGTFRDLAVVLETPALITALDIFVDRRYSALP 384
Db 190 MDMLKFLHGGGAEL--HSILSRVQDLGADTVQIVAIT-DRTKVLVHAIKOLKAAMLNAYP 246
```

```
QY 385 VV-----NECQVQVGLYSRFDV--IHLAAQOTYNNHLD-MSVGEALRQRTLC 426
Db 247 IVRATGVGQDDHKOLINGRCRKLIGTFSATDLRGCHTSSLSKWLGISALAFTEEVRSPL 306
QY 427 CLEG-----VLSQPHESLGEVIDRIAREQVHRLVLVDETOQHLLGVVSLSDILQ 475
Db 307 YSESDMQNREGSSRELVTCTAESPILSEVLEKAVTSHVHVWVVDQEGLLVGVVSLTDVIR 366
QY 476 AL 477
Db 367 VI 368

RESULT 9
T46197
hypothetical protein T8P19.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46197
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F..
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23008
A:Accession: T46197
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <CHO>
A:Cross-references: EMBL:AL133315
A:Experimental source: cultivar Columbia; BAC clone T8P19
C:Genetics:
A:Map position: 3
A:Introns: 69/1; 126/3; 178/3; 270/3; 325/2
A:Note: T8P19.40

Query Match 7.8%; Score 198; DB 2; Length 424;
Best Local Similarity 22.5%; Pred. No. 1.4e-07;
Matches 89; Conservative 82; Mismatches 140; Indels 84; Gaps 18;
QY 162 SPOAPPKLG-----WDELKPKGAQIYVREMOSHCTCYDAMATSSKLVIPTDM-LEIK- 213
Db 23 SPEA---KLGMRVEDLWDBQ--KPQLSPNEKL--NACFESIPVSAFPLSSDSQDIEIRS 74
QY 214 -----KAPFALVANGVRAAPLWD-----SKKQSFQVGLMTITDFTLV-VLHRY---RS 256
Db 75 DTSLAEAQVQTLKFKVLSAPVVDVADAPEDASWIDRYIGIVEFGIVVWLLHQLLEPPSPRS 134
QY 257 PLVQIYEIQH-----KIEWREIY-----LQGCFA--PLYSI 287
Db 135 PAVAASNGFSHDFTDVLNDGDSAVTSGNFFEVLTSSSELYKNTKVRDIDISTFRWAPFLAL 194
QY 288 SPNDSLFEAVYTLTKNRIHRLPVDLPVSGNVHLHILTHKRLLKFLHIFGSLPPSPFLYRT 347
Db 195 QKENSPLTMLLLSKYKMKSIIPVVDLGVAKIENIITQSGVIHMLAECAGLLWFEDWGIKT 254
QY 348 IQDLGIGTFR--DLAVVLETPALITALDIFVDRRYSALPVVNGCG--VVGLYSRFDV-IH 403
Db 255 LSEVGLPIMSDDHIKIYDEPVLQAFKLMRRKIGGIPVIERNSEKPVGNISLSDVQFL 314
QY 404 LAAQOTYN-----HLDMSVGEALRQRTLC-----LEGVLSQPHESLGEVIDRIA 448
Db 315 LTAPETIHYDRSYTTKNFLVSVREHLEK---CGDTSAPIMSGVIACKTNTKELLMLD 371
QY 449 REQVHRLVLVDETOQHLLGVVSLSDILQALVLSGAP 483
Db 372 AEKIHRIYVDDFGMLEGLITLRDIARLVHPEPSG 406

RESULT 10
T18227
hypothetical protein - yeast (Candida albicans)
C:Species: Candida albicans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18227
```



Db 195 NI--TGPELVESASAYAMLSDMLISFFDQSSQLHGLSHVTVDLSAIHNTVLAITSQAR 253  
QY 371 ---ALDIFDVRVSALPVMWECGQ-----VWGLSRFDV--IHLLAQOTVN 411  
Db 254 VKDAIOCSIAMLNAPVIVEAGEGEGHQLVDGKNRVRVGTFSASDLKCHLALRSW- 312  
QY 412 HLDMSVGEALRQ--RTLCLEG-----VLSQPHESGLGEVDRIAREQVHRLVLVDFT 461  
Db 313 -LPLNALEREVEKIPRTLFTAAVSTPGRELVTCHVTSTLAQVIMHVTTKRVHRVWVDQN 371  
QY 462 QHLLGVVSLSDILOALVLS 480  
Db 372 GGLQGLVSLTDIIF-AVWRS 389  
RESULT 13  
B96720  
hypothetical protein T17F3.17 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: B96720  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
C:Accession: B96720  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-447 <STO>  
A:Cross-references: GB:AE005173; NID:g6358813; PIDN:AAF07393.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T17F3.17  
A:Map position: 1  
Query Match 6.4%; Score 163; DB 2; Length 447;  
Best Local Similarity 25.5%; Pred. No. 8.1e-05;  
Matches 63; Conservative 39; Mismatches 107; Indels 38; Gaps 10;  
QY 262 YEI----EQHKIETWREIYLOGCFKPLVSLSPNDLSLFEAVYTLIKNRIHRLPVLDPVSCN 317  
Db 194 YEVILQEEFKSTTVRTILKSPRAWPELPVSTESSMLNVMLLSKYRLRNPVVKTCPEPD 253  
QY 318 VLHILTHKRLK-----FMIHTGSLRPRSPFLVITQDLGIGTFRDLAVV-LETA 366  
Db 254 IKNVVTSAAVHGLEGCKGRWFDHI--SALP-----ISDLGLPFMSPNEVIESIESE 303  
QY 367 P-ILTALDIFDVRVSALPVMWECG-QVGLYSRFDV-IHLLAQO-----TYNHLDM 415  
Db 304 ELILEAFKMRDNNITGLPVVVEGLNKRIVGNISMRDIRYLLIQPEVFSNFRQLTKVSFAT 363  
QY 416 SVGEALRQRTLCLEGVLSQPHESGLGEVDRIAREQVHRL-VLVDFTQHLLGVVSLSDIL 474  
Db 364 KITAGEEGLAIPAI-TCPDPSPLGVSINLSKRSVHRVVAAGDENELYGVITLROVI 422  
QY 475 QALVLSF 481  
Db 423 SCFVSEP 429  
RESULT 14  
G75056  
dehydrogenase PAB0961 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: G75056  
R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome :  
A:Reference number: A75001  
A:Accession: G75056  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-392 <RAW>  
A:Cross-references: GB:AJ48287; GB:AL096836; NID:g5458657; PIDN:CAB50348.1; PID:e15  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB0961  
Query Match 6.4%; Score 161.5; DB 2; Length 392;  
Best Local Similarity 23.4%; Pred. No. 8.8e-05;  
Matches 63; Conservative 54; Mismatches 103; Indels 49; Gaps 12;  
QY 227 APLWDSKKQSFVGMLTITDITLVL---HRYRSPVLQVIEIQHKIETW-----RELY 276  
Db 21 APLSEA-----IGIIEKEDPDLLIVFDNWKVGLVTDQLIIRSHL--KWDPTKAKYRVQY 73  
QY 277 LQGCFFPLVSIISPNDLSLFEAVYTLIKNRIHRLPVLDPVSGNVHLHTKRLKFLHIFGS 336  
Db 74 -----KPAPVVKPTDGLSHAALLLETDLRSLFVGEN-KAELGLVISDMALLE----- 120  
QY 337 LLPRPSFLYRTIQDLGIGTFRDLAVVLETAPIILTAIDIFDVRVSALPVMWECGQVVGLY 396  
Db 121 RVAAEEFGKRKVEEP---MTKDVITLGPDDTVAKALATMRDHGISRPVYVDEGKLEGIV 177  
QY 397 SRDVI-----HLAAQOTYNHLDMSVGE-----ALRQRTLCLEGVLSQPHESLGEVID 445  
Db 178 TLHDLIRITKPRFAQ--YGEI---AGEKIPFMSKUREAMKINGVITIMPEATIREAVS 232  
QY 446 RIAREQVHRLVLDVDETQHLGLGWVSLSDIL 474  
Db 233 TMRDNNIDGLVVDENKVKVGLITVKDLL 261  
RESULT 15  
S46088  
hypothetical protein YBR214w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YBR1501  
C:Species: Saccharomyces cerevisiae  
C:Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 19-Apr-2002  
C:Accession: S46088; S46090  
R:Rieger, M.  
submitted to the Protein Sequence Database, August 1994  
A:Reference number: S45734  
A:Accession: S46088  
A:Molecule type: DNA  
A:Residues: 1-527 <RIE>  
A:Cross-references: EMBL:Z36083; NID:g536599; PIDN:CAA85178.1; PID:g536600; MIPS:YBR:  
A:Experimental source: strain S288C  
R:Dubois, E.; El Bakoury, M.; Glausdorff, N.; Messenguy, F.; Pierard, A.; Scherens,  
submitted to the Protein Sequence Database, August 1994  
A:Reference number: S45782  
A:Accession: S46090  
A:Molecule type: DNA  
A:Residues: 114-527 <DUB>  
A:Cross-references: EMBL:Z36083; MIPS:YBR214w  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:SDS24  
A:Cross-references: SGD:S0000418  
A:Map position: 2R  
Query Match 6.2%; Score 158; DB 2; Length 527;  
Best Local Similarity 21.5%; Pred. No. 0.00025;  
Matches 94; Conservative 63; Mismatches 182; Indels 98; Gaps 17;  
QY 71 PPGCGECPSSRPAESTGL--EATFPKTTPLAQADPAG-----VGTPTGWDCLPSDCT 122  
Db 9 PPSOSNSSNNLPSTRSHASIVEMLSPTPLPLPHVQVNDTDDKEQPEESTPPT-----AT 60

```
QY 123 ASAAGSSTDDVELATEPATEAMECELEGLLEERPALCLS-----POAPEPKLG-----W 172
Db 61 AAAPGPGC---AATPAP-----LRDEKPFKLSAVPMTQTPSQCLSCVHAQKW 105
QY 173 DDELEKPGAQIYMFMEHCTCYDANATSSKLVIFDTWLEIKKAFPAFFALVANGYRAAPLWDS 232
Db 106 -----QHPLSQLIEONKLI FVPGSISVEEAFNTLIKYLHNSIPV--- 145
QY 233 KKOSFVGMLTITDF-----ILVLRHYRSPLVQIYEIEQHKIETWREIYLQGCCKP 283
Db 146 --ESPFGDMNCFTDYNDLNSYLLLVLNKITVSNKQLTADCQNGKPPVPGEMVKLTPKNP 203
QY 284 LVSI SPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLKPL----HIFGSLLP 339
Db 204 FYKL PENESL-STVMGILSGVHRVAITNEEMTKVKGILSORRLIKYLDNARSTSLP 262
QY 340 RPSFLYRTIQDLGIG-----TFRDLAV--VLETAPILTALDIFVDRRVSA LPV VNEC 389
Db 263 ---LLNSSQLDLHGVLNIQSKPTSHQSRVISIOGEEPLMGLYKMHVERISSIAVIDKQ 319
QY 390 GOVVGLYSRFDVIHLAAQQTYNHLDMS-----VGEALRQRTICLE----GVLSCQPHESLG 441
Db 320 GNLLGNISVTDVKHVRTSQYPLHKTKCHRFISVILNSRGLTGKDSFPFIEHVPSSSIA 379
QY 442 EVIDRIAREQVHRLVIV 458
Db 380 RTLAKLVATKSHRLWIV 396
```

Search completed: June 6, 2003, 11:03:34  
Job time : 34.6865 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 17.9842 Seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: US-09-826-581-6

Perfect score: 2538

Sequence: 1 MEPLGHEALRRTPSNSSLG.....ISDILQALVSPAGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	94.5	464	1 AAKI_HUMAN	Q9ug19 homo sapien
2	2011	79.2	464	1 AAKI_PIG	Q9m9p4 sus scrofa
3	1093	43.1	569	1 AAKH_HUMAN	Q9ugj0 homo sapien
4	1028	40.5	330	1 AAKG_RAT	P80385 rattus norv
5	1017	40.1	331	1 AAKG_HUMAN	P54619 homo sapien
6	1013	39.9	330	1 AAKG_BOVIN	P58108 bos taurus
7	997	39.3	330	1 AAKG_MOUSE	O54950 mus musculu
8	457	18.0	322	1 SNF4_YEAST	P12904 saccharomyc
9	456.5	18.0	334	1 YDRI_SCHPO	Q10343 schizosacch
10	418.5	16.5	328	1 SNF4_KLULA	Q9p869 kluyveromyc
11	267.5	10.5	133	1 AAKG_PIG	Q09138 sus scrofa
12	158	6.2	527	1 YB64_YEAST	P38314 saccharomyc
13	149	5.9	280	1 YC25_METJA	Q58622 methanococc
14	139	5.5	300	1 YB33_THERPE	P15889 thermofilum
15	138.5	5.5	421	1 Y604_METJA	Q58799 methanococc
16	138	5.4	485	1 IMDH_PYRFU	P42851 pyrococcus
17	136.5	5.0	527	1 YGf6_METJA	Q58172 saccharomyc
18	126.5	5.0	168	1 YG26_METJA	Q58821 methanococc
19	126	5.0	486	1 IMDH_PYRHO	O58045 pyrococcus
20	119	4.7	485	1 IMDH_PYRAB	Q9uy49 pyrococcus
21	114	4.5	1053	1 CAPP_SYNPG	P06516 synechococc
22	113.5	4.5	296	1 YC32_METJA	Q58629 methanococc
23	110	4.3	496	1 IMDH_METJA	Q59011 methanococc
24	108.5	4.3	944	1 VGLB_HSVT2	Q9wr15 herpesvirus
25	107.5	4.2	509	1 Y100_METJA	O57564 methanococc
26	104.5	4.1	196	1 Y325_METKA	P50100 methanopyru
27	104	4.1	329	1 Y529_HELPY	Q25971 helicobacte
28	103	4.1	537	1 IMDH_DROME	Q07152 drosophila
29	102.5	4.0	2377	1 CCAG_HUMAN	O43497 homo sapien
30	102	4.0	138	1 Y922_METJA	Q58332 methanococc
31	101.5	4.0	424	1 Y146_MYCGE	Q49399 mycoplasma
32	101.5	4.0	835	1 U152_HSVSA	P14346 herpesvirus
33	101	4.0	521	1 IMDH_CHLVI	O50316 chlorobium

#### RESULT 1

ID	AAKI_HUMAN	STANDARD;	PRT;	464 AA.
AC	Q9UG19; Q9NRL1;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)			
DE	(AMPK gamma3).			
GN	PRKAG3 OR AMPKG3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20164049; PubMed=10698692;			
RA	Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;			
RT	"Characterization of AMP-activated protein kinase gamma-subunit			
RL	isoforms and their role in AMP binding.";			
RN	Biochem. J. 346:659-669(2000).			
RC	SEQUENCE FROM N.A.			
PC	TISSUE=Skeletal muscle;			
RX	MEDLINE=20280150; PubMed=10818001;			
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,			
RA	Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,			
RA	Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,			
RA	Andersson L.;			
RT	"A mutation in PRKAG3 associated with excess glycogen content in pig			
RT	skeletal muscle.";			
RL	Science 288:1248-1251(2000).			
CC	-!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID			
CC	SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO			
CC	REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND			
CC	INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-			
CC	SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT IT MAY PLAY A ROLE			
CC	IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.			
CC	-!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A			
CC	GAMMA NON-CATALYTIC REGULATORY SUBUNITS.			
CC	-!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART			
CC	AND PANCREAS.			
CC	-!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA			
CC	SUBUNIT FAMILY.			
CC	-!- SIMILARITY: CONTAINS 4 CBS DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AJ249977; CAB65117.1; ALT_INIT.			
DR	EMBL; AF214519; AAF73987.1; -.			

```
DR Genew; HGNC:9387; PRKAG3.
DR MM; 604976; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 147 201 CBS 1.
FT DOMAIN 228 282 CBS 2.
FT DOMAIN 303 356 CBS 3.
FT DOMAIN 375 428 CBS 4.
FT CONFLICT 58 58 T -> A (IN REF. 1).
FT CONFLICT 163 164 MQ -> JE (IN REF. 1).
FT CONFLICT 398 398 Q -> K (IN REF. 1).
FT CONFLICT 461 464 ALGA -> PSGPKI (IN REF. 1).
SQ SEQUENCE 464 AA; 51514 MW; 53985C2C77003A63 CRC64;

Query Match
Best Local Similarity 94.5%; Score 2398; DB 1; Length 464;
Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 26 MSFLEQENSSWSPAVTSSSIRIGKRRKALKRWTRQKSVEEGPPQGEGPRSPAAE 85
Db 1 MSFLEQENSSWSPAVTSSSIRIGKRRKALKRWTRQKSVEEGPPQGEGPRSPAAE 60

Qy 86 STGLEATFPKTTPLAQADPAGVPTTGWDCPLSDCTASAAGSSDDVELATEFPATEAW 145
Db 61 STGLEATFPKTTPLAQADPAGVPTTGWDCPLSDCTASAAGSSDDVELATEFPATEAW 120

Qy 146 ECELEGLEERPALCLSPQAPFPKLGWDDLRKPGAIYMRFMQBHCTCYDAMATSSKLVI 205
Db 121 ECELEGLEERPALCLSPQAPFPKLGWDDLRKPGAIYMRFMQBHCTCYDAMATSSKLVI 180

Qy 206 FDTMLETKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPLVQIYEIE 265
Db 181 FDTMLETKKAFFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPLVQIYEIE 240

Qy 266 QHKIETWREIYLOGCFKPLVSLSPNDSLFEAVYTLKKNRHLRPLVDPVSGNVHLIHTK 325
Db 241 QHKIETWREIYLOGCFKPLVSLSPNDSLFEAVYTLKKNRHLRPLVDPVSGNVHLIHTK 300

Qy 326 RLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFDVRVSLPV 385
Db 301 RLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFDVRVSLPV 360

Qy 386 VNECCQVVGLYSRFDVIHLAAQQTYNHLDMSVGEALKQRTCLLEGVLSQCPHESLGEVID 445
Db 361 VNECCQVVGLYSRFDVIHLAAQQTYNHLDMSVGEALKQRTCLLEGVLSQCPHESLGEVID 420

Qy 446 RIAREQVHRLVLDVETQHLLGVVSLSDILQALVLSPAGIDALGA 489
Db 421 RIAREQVHRLVLDVETQHLLGVVSLSDILQALVLSPAGIDALGA 464

RESULT 2
AAKL_PIG
ID AAKL_PIG STANDARD; PRT: 464 AA.
AC Q9MYT4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)
DE (AMPK gamma3).
GN PRKAG3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A., AND VARIANT RN(-) GLN-200.
RC RC
RP TISSUE-Skeletal muscle;
RX MEDLINE=20280150; PubMed=10818001;
RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,
RA Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,
```

```
RA Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,
RA Andersson L.;
RT "A mutation in PRKAG3 associated with excess glycogen content in pig
RT skeletal muscle.";
RL Science 288:1248-1251(2000).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPIASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE
CC IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -!- TISSUE SPECIFICITY: MUSCLE.
CC -!- DISEASE: DEFECTS IN PRKAG3 (RN-) ARE ASSOCIATED WITH EXCESS
CC GLYCOGEN CONTENT (ABOUT 70%) IN SKELETAL MUSCLE. THIS MUTATION
CC ORIGINATED IN THE HAMPSHIRE BREED PIGS AND HAS BENEFICIAL EFFECTS
CC ON MEAT CONTENT BUT DETRIMENTAL EFFECTS ON PROCESSING YIELD. THUS,
CC THIS MUTATION IS OF CONSIDERABLE ECONOMIC SIGNIFICANCE IN THE PIG
CC BREEDING INDUSTRY.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF214521; AAF73989.1; -.
DR EMBL; AF214520; AAF73988.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation.
FT DOMAIN 172 226 CBS 1.
FT DOMAIN 253 307 CBS 2.
FT DOMAIN 328 381 CBS 3.
FT DOMAIN 400 453 CBS 4.
FT VARIANT 200 200 R -> Q (IN RN-).
SQ SEQUENCE 464 AA; 51308 MW; 17638CB12A2BA9DF CRC64;

Query Match
Best Local Similarity 79.2%; Score 2011; DB 1; Length 464;
Matches 400; Conservative 19; Mismatches 44; Indels 2; Gaps 2;

Qy 26 MSFLEQENSSWSPAVTSSSIRIGKRRKALKRWTRQKSVEEGPPQGEGPRSPAAE 85
Db 1 MSFLEQENSSWSPAVTSSSIRIGKRRKALKRWTRQKSVEEGPPQGEGPRSPAAE 60

Qy 86 STGLEATFPKTTPLAQADP-AGVGTPTTGWDCPLSDCTASAAGSSDDVELATEFPATEA 144
Db 61 STGLEATFPKTTPLAQADP-AGVGTPTTGWDCPLSDCTASAAGSSDDVELATEFPATEA 120

Qy 145 WECELEGLEERPALCLSPQAPFPKLGWDDLRKPGAIYMRFMQBHCTCYDAMATSSKLVI 204
Db 121 SODEL-GLVEEKAPCPSPVLLPRLGWDDLRKPGAIYMRFMQBHCTCYDAMATSSKLVI 179

Qy 205 IFDNLMEIKKAFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPLVQIYEI 264
Db 180 IFDNLMEIKKAFALVANGVRAAPLWDSKQSFVGMLTITDFILVLRHYRSPLVQIYEI 239

Qy 265 EQHKIETWREIYLOGCFKPLVSLSPNDSLFEAVYTLKKNRHLRPLVDPVSGNVHLIHT 324
Db 240 EQHKIETWREIYLOGCFKPLVSLSPNDSLFEAVYTLKKNRHLRPLVDPVSGNVHLIHT 299

Qy 325 RLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFDVRVSLPV 384
Db 300 RLLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFDVRVSLPV 359
```





```

DR MM; 602743; -
DR MM; 194200; -
DR DR InterPro: IPR000644; CBS_domain.
DR PFam: PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
DR PAM; 194200; -
KW Fatty acid biosynthesis; Repeat; CBS domain; Alternative splicing;
KW Disease mutation..
FT DOMAIN 274 328 CBS 1.
FT DOMAIN 355 409 CBS 2.
FT DOMAIN 430 483 CBS 3.
FT DOMAIN 502 555 CBS 4.
FT VARSPLIC 1 241 MISSING (IN ISOFORM B).
FT VARIANT 302 302 R -> Q (IN WPWS).
FT VARIANT 350 350 /FTIG-VAR_013264.
FT VARIANT 383 383 SEVERE).
FT VARIANT 383 383 /FTIG-VAR_013265.
FT VARIANT 400 400 H -> R (IN FAMILIAL HCM WITH WPWS;
FT VARIANT 400 400 SEVERE).
FT VARIANT 400 400 /FTIG-VAR_013266.
FT VARIANT 400 400 T -> N (IN WPWS WITH CARDIAC
FT VARIANT 400 400 HYPERTROPHY).
FT VARIANT 488 488 /FTIG-VAR_013267.
FT VARIANT 488 488 N -> I (IN WPWS WITH CARDIAC
FT VARIANT 531 531 HYPERTROPHY).
FT VARIANT 531 531 /FTIG-VAR_013268.
FT VARIANT 531 531 R -> G (IN WPWS; ABSENCE OF CARDIAC
FT VARIANT 531 531 HYPERTROPHY; ONSET IN CHILDHOOD).
FT VARIANT 531 531 /FTIG-VAR_013269.
FT CONFLICT 218 233 ASPTHVAPSKAAALAA -> MLIAVLLPLRWR
FT CONFLICT 218 233 (IN REF 5).
FT SEQUENCE 569 AA; 63066 MW; F51C30668C294089 CRC64;
Query Match 43.1%; Score 1093; DB 1; Length 569;
Best Local Similarity 51.0%; Pred. No. 3.2e-75;
Matches 231; Conservative 75; Mismatches 95; Indels 52; Gaps 10;
QY 63 QKVEEKEP---PCQEGE-----PRSR-----PAESTGLEATFPKTPPLA--QADP--- 104
DQ 128 RSKSKESPNPATSPGIRFBSRSRKTGLSSSPSTPQTKQTFPLESYKHEPERL 187
QY 105 -----AGVPTPTGMDCLPSDC---TASAGSST-----DDVELATEFPATEMECELEG 151
DQ 188 ENRIYASSPPDGTGRFCPSQSPTRPPLASPHYAPSKAAALAAALGAPAA-----G 241
QY 152 LLEERPALCLSPQAPFKIGWDE--LRKPCAQIYMRPMQBHTCYDAMATSSKLVFTDTL 210
DQ 242 MLE-----KLEFDEAVEDESEGVYMRPMRSHKCYDIPVTSKLVVFTTL 287
QY 211 EIKKAFALVANGVRAAPIMUSKKOSPVGMLTTFDFILVILHRYRSPVQIYETEQHKIE 270
DQ 288 QVKAFTALVANGVRAAPIMUSKKOSPVGMLTTFDFILVILHRYRSPVQIYETEQHKIE 347
QY 271 TWREIYLQGCFKPLVLSIPNDISIPEAVVTLIKNN[HLRPLVDPVSGVNLHLTHKRLKF 330
DQ 348 TWREIYLQGCFKPLVLSIPNDISIPEAVVTLIKNN[HLRPLVDPVSGVNLHLTHKRLKF 407
QY 331 LHIFGSLPRPFLYRTIOLGIGTRDLAVVLETPALITFALDFVDRVSPALPVNVEG 390
DQ 408 LQLFMDMPKPAFMKQNLDELIGTYHNIATFHPDTPIKALNIFVERKISALPVNVEG 467
QY 391 QVVGILSRFDVILHAQQIYNHLDMSVGEALRQKTLCEGLVSCQPHSLEGEVLDIARE 450
DQ 468 KVDIYKSFVDVILHAQQIYNHLDMSVGEALRQKTLCEGLVSCQPHSLEGEVLDIARE 527
QY 451 QVHRLVLVDETHLLGWLSLSLILQALVLSAG 483
DQ 528 EVHRLVNVNDEADSVGLSLISLILQALVLSAG 560
RESULT 4
AAKG_RAT
ID AAKG_RAT STANDARD; PRG; 330 AA.

```

```

AC P80385;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
DE PKAG1.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=96215327; PubMed=8626596;
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,
RA Beri R.K., Carling D.;
RT *Characterization of AMP-activated protein kinase beta and gamma
RT subunits. Assembly of the heterotrimeric complex in vitro.*;
RL J. Biol. Chem. 271:10282-10290(1996).
RN [2]
RP SEQUENCE OF 8-330 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT *Non-catalytic beta- and gamma-subunit isoforms of the
RT 5'-AMP-activated protein kinase.*;
RL J. Biol. Chem. 271:8675-8681(1996).
RN [3]
RP SEQUENCE OF 48-330 FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;
RT *Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
RT homologs of proteins that interact with yeast Snf1 protein kinase.*;
RL J. Biol. Chem. 269:29343-29346(1994).
CC -I- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HORMONE-SENSITIVE LIPASE AND
CC HYDROXYMETHYLGUTARYL-COA REDUCTASE. APPEARS TO ACT AS A METABOLIC
CC STRESS-SENSING PROTEIN KINASE SWITCHING OFF BIOSYNTHETIC PATHWAYS
CC WHEN CELLULAR ATP LEVELS ARE DEPLETED AND WHEN 5'-AMP RISES IN
CC RESPONSE TO FUEL LIMITATION AND/OR HYPOXIA. THIS IS A REGULATORY
CC SUBUNIT.
CC -I- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT; A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART AND BRAIN, ALSO
CC FOUND IN KIDNEY, WHITE ADIPOSE TISSUE, LUNG AND SPLEEN.
CC -I- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -I- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; X95578; AAC64831.1; -.
CC EMBL; U42413; AAC52580.1; -.
CC InterPro: IPR000644; CBS_domain.
CC Pfam; PF00571; CBS; 4.
CC SMART; SM00116; CBS; 4.
CC Fatty acid biosynthesis; Repeat; CBS domain.
CC DOMAIN 47 95 CBS 1.
CC DOMAIN 122 176 CBS 2.
CC DOMAIN 196 249 CBS 3.
CC DOMAIN 270 322 CBS 4.

```





```

RL Gene 67:247-257(1988).
RP [2]
RX SEQUENCE FROM N.A.
RA MEDLINE=90097921; PubMed=2481228;
RA Celenza J.L., Eng F.J., Carlson M.;
RT "Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae:
RT evidence for physical association of the SNF4 protein with the SNF1
RT protein kinase.";
RL Mol. Cell. Biol. 9:5045-5054(1989).
RP [3]
RX SEQUENCE FROM N.A.
RA Lauquin G.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RP [4]
RX SEQUENCE OF 1-21 FROM N.A.
RA Doi A., Doi K.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RP [5]
RX SEQUENCE OF 30-34 AND 316-322.
RA MEDLINE=94131988; PubMed=7905477;
RA Mitchellhill K.I., Stapleton D., Gao G., House C., Michell B.,
RA Katsis F., Witters L.A., Kemp B.E.;
RT "Mammalian AMP-activated protein kinase shares structural and
RT functional homology with the catalytic domain of yeast Snf1 protein
RT kinase.";
RL J. Biol. Chem. 269:2361-2364(1994).
CC -!- FUNCTION: THIS PROTEIN CAUSES EXPRESSION OF GLUCOSE-REPRESSIBLE
CC GENES UPON GLUCOSE DEPRIVATION. IT INTERACTS AND HAS FUNCTIONAL
CC RELATIONSHIP TO THE PROTEIN-KINASE SNF1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21760; AAA34472.1; -
DR EMBL; M30470; AAA35061.1; -
DR EMBL; Z72637; CAA96823.1; -
DR EMBL; D16506; BAA03958.1; -
DR PIR; JT0316; RBYC3.
DR PIR; A33480; A33480.
DR SGD; S0003083; SNF4.
DR InterPro; IPR000544; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Carbohydrate metabolism; Transcription regulation; Nuclear protein;
KW Repeat; CBS domain.
FT DOMAIN 35 89 CBS 1.
FT DOMAIN 117 175 CBS 2.
FT DOMAIN 192 246 CBS 3.
FT DOMAIN 259 318 CBS 4.
SQ SEQUENCE 322 AA; 36401 MW; 51B387E346EE9561 CRC64;
Query Match 18.0%; Score 457; DB 1; Length 322;
Best Local Similarity 34.2%; Pred. No. 2e-27;
Matches 106; Conservative 66; Mismatches 108; Indels 30; Gaps 6;
OY 186 RMQWHTCYDAMATSSKLVTFDMLLEIKKAFALVANGVRAAPLWDSKKQSFVGLMTITD 245
DB 24 KFNLSKTSYDVLPSYSLVLDLTSLLVKKASLVNQLQNSIVSAPLWDSKTSRFAGLLTTTD 83
OY 246 FILVLRHYRGPVLPVQVIEIQHXIEWRREIVLQGCXPL-----VSISPNDSLF 294
DB 84 FINVQYFESNP-----DFELVDKLQLDG-LKDIERALGVQDLDTASIHFSRPLF 133
OY 295 EAVYTLIKNRHLRFLVDPVSGN-----VLHLTHKRLKLFHIFGSLLPSPS-FLYRTIQ 349

```

```

DB 134 EACLKWLSSGRPLIDQDEETHREIVSVLTQYRILKFV-----ALNCRETHLKPIG 189
OY 350 DLGIGTFRDLAWVLETAPILTALDIFVDRVSALPVNCCQVGVLYRFDVIHLAAQOT 409
DB 190 DLNITQDNMSQMTTPVIVDTQMLTOGRVSSVPIDENGYLINVYAYDVLGIKGGI 249
OY 410 YNHLDSVGEALRQTLCLGVLSCOPHESIGCEVDTRIAREOVHRLVLDVDETHLLGVVS 469
DB 250 YNDLSVSGEALMRSSDDFEGVYCTCKNDKLTINDNIRKARHREFVDDVGRVLGVLT 309
OY 470 LSDILQALVL 479
DB 310 LSDILKYILL 319
RESULT 9
YDAL_SCHPO
ID YDAL_SCHPO STANDARD; PRT; 334 AA.
AC Q10343; Q9UTJ1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Cif12.0ic in chromosome I.
GN SPAC1F12.0IC OR SPAC1556.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=2184801; PubMed=11859360;
RA Wood V., William R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skeilton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volktaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69944; CAA93805.1; -

```





```
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Goughen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67563; AB999228.1; -.
DR TIGR: MJ1225; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain; Complete proteome.
FT DOMAIN 7 62 CBS 1.
FT DOMAIN 88 141 CBS 2.
FT DOMAIN 152 204 CBS 3.
FT DOMAIN 227 280 CBS 4.
SQ SEQUENCE 280 AA; 31719 MW; 614AE160C92E45F CRC64;

Query Match 5.9%; Score 149; DB 1; Length 280;
Best Local Similarity 19.6%; Pred. No. 0.00035;
Matches 61; Conservative 72; Mismatches 126; Indels 52; Gaps 11;

QY 183 IYRPMQEHCTYDAMATSSKLVIFDPMLEIKKAFALVANGVRAAPLWDSKQSPVGMLT 242
DB 1 MEVRVYMK-----IAQNKRVIVPTTIRKALMTWNEKNRYRLPVPVWAGNKKYVGII 53
QY 243 ---IPDFT-----LVHRYRSPVLQVIEIOHKETWREIYLGQCFKPLVSPNDS 292
DB 54 SMDIVDFMGSGSKYNLIREKHENFLAINEPVREIME-----ENVITLKENAD 102
QY 293 LFEAVYTLIKNRIHRLPVLDPVSGVNLHILTHKRLKFLHFGSLPRSPFLYRTIQDLG 352
DB 103 IDEALTEFLTNKGGAPIVND-FNQILSLITERDQVIRAL-----LDKIDENEV 149
QY 353 IGTFF--RDLAVVLETPALILTALDILVDRRVSAALPVVNECGVGLYSRFDVTH-LAAQOT 409
DB 150 IDYITRDIIVATPGEKLDKDVARTWVRNGFRRLPVSE-GRLVGIITSTDFIKLGSQWA 208
QY 410 YNHLDMVSVEA-----LRQITCLGVLSCQPHESLGEVIDRIAREQVHRLVLDVTOHL 464
DB 209 FNH--MQTGNVREITNVREINMKRDVITAKEGDKLAKIAEIMVNDIGALPVDENLRI 266
QY 465 LQVLSLSILQ 475
DB 267 KGIITEKDVLK 277

RESULT 14
YR33_THEPE
ID YR33_THEPE STANDARD; PRT; 300 AA.
AC P15889;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 33.4 kDa protein in ribosomal RNA operon.
OS Thermophilum pendens.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermofilaceae; Thermofilum.

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Goughen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14835; CA832944.1; -.
DR PIR: S08244; S08244.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 8 61 CBS 1.
FT DOMAIN 87 140 CBS 2.
FT DOMAIN 150 202 CBS 3.
FT DOMAIN 224 276 CBS 4.
SQ SEQUENCE 300 AA; 33437 MW; D811A313D37A4293 CRC64;

Query Match 5.5%; Score 139; DB 1; Length 300;
Best Local Similarity 23.4%; Pred. No. 0.0022;
Matches 47; Conservative 52; Mismatches 86; Indels 16; Gaps 6;

QY 279 GCFKPLVSPNDSLSLFEAVYTLIKNRIHRLPVLDPVSGVNLHILTHKRLKFL--HIFGS 336
DB 9 GREPLAVVPSSRLVDVLVANGRNVRVPLVDE-RGVLGKGVARSADLVDFLGGRRFRD 67
QY 337 LIPR--PSFLYRTIQDLGIGTFR-DLAVVLETPALILTALDIFVDRRVSAALPVVNECGV 393
DB 68 VVEARFNGVDYKALEQGTGVFEKYDPPYVYTRSDLRVIELMVERGIGALAVVDELRV 127
QY 394 GLYSRFDVHIAAQOTTNHLDMVSGEALQRITCLGVLSCQPHESLGEVIDRIAREQVH 453
DB 128 GIVSERHVVSLA-----NVETHVKVKEIMTSEVYVLSPMDSLFEGMRVMSERRIR 178
QY 454 RLVLVDETOHLLGVVLSLSDIL 474
DB 179 RLPLV-SGEELRGIVTIKDVL 198

RESULT 15
YE04_METJA
ID YE04_METJA STANDARD; PRT; 421 AA.
AC Q58799;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1404.
GN MJ1404.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
```

```
OX NCBI_TaxID=2269;
RW SEQUENCE FROM N.A.
RC STRAIN=HWV3 / DSM 2475;
RA Klems J., Jeffers H., Olesen T., Ingelore H., Garrett R.A.;
RT "Sequence, organisation and transcription of the ribosomal RNA operon
RT and the downstream tRNA and protein genes in the archaeobacterium
RT Thermofilum pendens.";
RL Syst. Appl. Microbiol. 13:117-127(1990).
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14835; CA832944.1; -.
DR PIR: S08244; S08244.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 8 61 CBS 1.
FT DOMAIN 87 140 CBS 2.
FT DOMAIN 150 202 CBS 3.
FT DOMAIN 224 276 CBS 4.
SQ SEQUENCE 300 AA; 33437 MW; D811A313D37A4293 CRC64;

Query Match 5.5%; Score 139; DB 1; Length 300;
Best Local Similarity 23.4%; Pred. No. 0.0022;
Matches 47; Conservative 52; Mismatches 86; Indels 16; Gaps 6;

QY 279 GCFKPLVSPNDSLSLFEAVYTLIKNRIHRLPVLDPVSGVNLHILTHKRLKFL--HIFGS 336
DB 9 GREPLAVVPSSRLVDVLVANGRNVRVPLVDE-RGVLGKGVARSADLVDFLGGRRFRD 67
QY 337 LIPR--PSFLYRTIQDLGIGTFR-DLAVVLETPALILTALDIFVDRRVSAALPVVNECGV 393
DB 68 VVEARFNGVDYKALEQGTGVFEKYDPPYVYTRSDLRVIELMVERGIGALAVVDELRV 127
QY 394 GLYSRFDVHIAAQOTTNHLDMVSGEALQRITCLGVLSCQPHESLGEVIDRIAREQVH 453
DB 128 GIVSERHVVSLA-----NVETHVKVKEIMTSEVYVLSPMDSLFEGMRVMSERRIR 178
QY 454 RLVLVDETOHLLGVVLSLSDIL 474
DB 179 RLPLV-SGEELRGIVTIKDVL 198

RESULT 15
YE04_METJA
ID YE04_METJA STANDARD; PRT; 421 AA.
AC Q58799;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1404.
GN MJ1404.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
```

Search completed: June 6, 2003, 11:01:28  
Job time : 19.9842 secs

Search completed: June 6, 2003, 11:01:28  
Job time : 19.9842 secs



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1085.5	42.8	566	11	Q91WC5		Q91wc5 mus musculus
2	871.5	34.3	647	5	Q96613		Q96613 drosophila
3	869.5	34.3	634	5	Q9VDD2		Q9vdd2 drosophila
4	869.5	34.3	906	5	Q8SX78		Q8sxt8 drosophila
5	869.5	34.3	1400	5	Q8SZS7		Q8szs7 drosophila
6	736.5	29.0	372	5	Q9BHL6		Q9bhl6 caenorhabdi
7	559.5	22.0	577	5	Q8SSV7		Q8ssv7 dictyosteli
8	488	19.2	423	5	Q92168		Q92168 caenorhabdi
9	470.5	18.5	448	5	Q9N501		Q9n501 caenorhabdi
10	374.5	14.8	487	10	Q944A6		Q944a6 arabidopsis
11	364	14.3	382	10	Q9FV59		Q9fv59 arabidopsis
12	348.5	13.7	497	10	Q9FUY5		Q9fuy5 zea mays (m
13	345	13.6	496	10	Q9FYU4		Q9fyu4 zea mays (m
14	243.5	9.6	391	10	Q04028		Q04028 arabidopsis
15	227	8.9	423	5	Q95ZL0		Q95z10 caenorhabdi
16	225	8.9	63	11	Q925V0		Q925v0 mus musculus

Db 218 SPTYHAPLRTAVLAAAPGPAE-AGMLE-----KLEFQEE-EDSESGFYMR 261

17	217.5	8.6	5	Q22022	Q22022 caenorhabdi
18	214	8.4	379	10	Q41108 phaseolus v
19	198	7.8	424	10	Q9s7w6 arabidopsis
20	196	7.7	629	3	Q94035 candida alb
21	194	7.6	435	10	Q8rz19 oryza sativ
22	190	7.5	373	10	Q9M727 lycopersico
23	189.5	7.5	352	10	Q9x137 arabidopsis
24	179	7.1	399	10	Q9SD00 arabidopsis
25	178.5	7.0	450	10	Q94D32 oryza sativ
26	174.5	6.9	278	17	Q8txy3 methanopyru
27	163	6.4	447	10	Q9CA83 arabidopsis
28	162.5	6.4	304	5	Q8t849 dictyosteli
29	161.5	6.4	392	17	Q9UYR4 pyrococcus
30	158.5	6.2	282	17	Q8ZVB6 pyrobaculum
31	158	6.2	396	3	Q8TFK4 kftzka
32	155.5	6.1	392	17	Q58317 pyrococcus
33	154	6.1	248	10	Q93XY6 arabidopsis
34	150.5	5.9	274	17	Q97314 sulfobus
35	150	5.9	392	17	Q8U3Y0 pyrococcus
36	148.5	5.9	250	17	Q97V95 sulfobus
37	146.5	5.8	335	5	Q8TFX0 dictyosteli
38	146	5.8	313	17	Q27292 methanobact
39	141.5	5.6	157	17	Q26740 methanobact
40	141.5	5.6	393	17	Q8T250 methanopyru
41	140.5	5.5	375	17	Q96V62 sulfobus
42	139	5.5	152	16	Q8YQ10 anabaena sp
43	137	5.4	284	17	Q29410 archaeoglob
44	135.5	5.3	501	17	Q8TWX4 methanopyru
45	134	5.3	355	17	Q97AJ9 thermoplasm

Qy 188 MQEHTCYDAMATSSKLVITDMLTKKAFFALVANGVRAAPLWDSKKQSFVGMLTITDFI 247  
Db 262 MRSKHCXDIIVPTSSKLVITDMLTKKAFFALVANGVRAAPLWESKKQSFVGMLTITDFI 321  
Qy 248 LVLRHYRSPVQVIEHQKLETRIEYLOGCEKPLVLSPNDSLFVAVYTLTKNLRHR 307  
Db 322 NLHRHYKSPVQVIEEHKEKLETRIEYLOETFKPLVNSPDASFDAVSLIKNKRHR 381  
Qy 308 LPVLDPVSGNVHLHTRKLLKFLHIFGSLLPRESFLYRTIODLGIGTFRDLAVVLETAP 367  
Db 382 LPVIDPISGNALYILTHKRLKFLQLFMSDKPAPKQNLDELGIGTYHNTAFTHPDP 441  
Qy 368 ILTALDIPDVRVSALPVVNCQVGVLYSRFDVHILAAQOTYHNLDMVSGALRQRTLC 427  
Db 442 IIKALNIFEVERISALPVVDSCKVVDIYSKFDVINLAEKTYNNIDITVTQALQHSQY 501  
Qy 428 LEGVLSQCPRESIGEVIDRTAREQVHRLVLVDEQHLGLGVVSLSDTLQALVLSAPG 483  
Db 502 FEGVVKCSKLETLTIVDRVRAEVHRLVVVNEADSIVGIIISUSDILQALILTPAG 557  
RESULT 2  
096613 ID O96613 PRELIMINARY; PRT: 647 AA.  
AC O96613;  
DT 01-MAY-1999 (TremBrel. 10, Created)  
DT 01-MAY-1999 (TremBrel. 10, Last sequence update)  
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)  
DE SNF4/AMP-activated protein kinase gamma subunit.  
GN SNF4A-GAMMA OR CG17299.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
CX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yoshida E.N., Benkel B.F., Fong Y., Hickey D.A.;  
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF094764; AAC95306.1; -  
DR EMBL; AF094763; AAC95305.1; -  
DR FlyBase; FBgn025803; SNF4A-gamma.  
DR InterPro; IPR000644; CBS\_domain.  
DR Pfam; PF00571; CBS; 4.  
DR SMART; SM00116; CBS; 4.  
DR KInase.  
SQ SEQUENCE 647 AA; 71592 MW; 8792BE1089730B52 CRC64;  
Query Match 34.3%; Score 871.5; DB 5; Length 647;  
Best Local Similarity 53.1%; Pred. No. 2.2e-62;  
Matches 164; Conservative 74; Mismatches 70; Indels 1; Gaps 1;  
Qy 175 ELRKPQAIYMRFMQEHCTCYDAMATSSKLVITDMLTKKAFFALVANGVRAAPLWDSKK 234  
Db 162 DLEEDDSQIFVKFRHKCYDLIPTSAKLVVFDQLLVKAFYALVYNGVRAAPLWDSK 221  
Qy 235 QSFVGMLTITDFILVLRHYRSPVQVIEHQKLETRIEYLOGCEKPLVLSPNDSLF 294  
Db 222 QQFVGMLTITDFIKILQYMKYKSPNASMEQLSEHKLDTWRDV-LHNQVMPVLSIGPDASLY 280  
Qy 295 EAVYTLTKNLRHLPVLDVPSGNVLHILTHKRLKFLHIFGSLLPRESFLYRTIODLGIG 354  
Db 281 DAKTLHRSRHLRPVLDIPDGNVLYILTHKRLRFLFLYINELPKPAQMKQSRLEKIG 340  
Qy 355 TFRDLAVVLETAPITALTALDIFVDRVSALPVVNCQVGVLYSRFDVHILAAQOTYHNLH 414  
Db 341 TYNNETADETTSITALKKFFVRRVSALPLVDSGRLVDIYAKFDVINLAEKTYNDLD 400  
Qy 415 NSVGALRQRTLCLEGVLSQCPRESIGEVIDRTAREQVHRLVLVDEQHLGLGVVSLSDIL 474  
Db 401 VSLRKANHRNENFEGVQKCNLDESILYTIMERIVRAEVHRLVVVNDENRKRVIIGIISLSDIL 460  
Qy 475 QALVLSAPG 483

Db 461 LXLVLRPSG 469  
RESULT 3  
Q9VDD2 ID Q9VDD2 PRELIMINARY; PRT: 634 AA.  
AC Q9VDD2;  
DT 01-MAY-2000 (TremBrel. 13, Created)  
DT 01-MAY-2000 (TremBrel. 13, Last sequence update)  
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)  
DE SNF4GAMMA protein.  
GN SNF4A-GAMMA OR SNF4GAMMA OR CG17299.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
CX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=BERKELEY;  
RA MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Branton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Turner R., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003733; AAF55864.1; -  
DR FlyBase; FBgn0025803; SNF4A-gamma.  
DR InterPro; IPR000644; CBS\_domain.  
DR Pfam; PF00571; CBS; 4.  
DR SMART; SM00116; CBS; 4.  
SQ SEQUENCE 634 AA; 70174 MW; 3FDD0D5E54BBE7C CRC64;  
Query Match 34.3%; Score 869.5; DB 5; Length 634;  
Best Local Similarity 53.1%; Pred. No. 3.1e-62;  
Matches 164; Conservative 73; Mismatches 71; Indels 1; Gaps 1;  
Qy 175 ELRKPQAIYMRFMQEHCTCYDAMATSSKLVITDMLTKKAFFALVANGVRAAPLWDSKK 234  
Db 175 ELRKPQAIYMRFMQEHCTCYDAMATSSKLVITDMLTKKAFFALVANGVRAAPLWDSKK 234

Db 149 DLEEDDSQIFVKFFRHKCYDLIPTSAKLIVFDTQLLVKAFYALVYNGVRAAPLWDSK 208  
Qy 235 QSFVGMLTITDFILVLRHYRSPVQIYEIQHKTWREIYLGQCFKPLVSI SPNDSLF 294  
Db 209 QQFVGMLTITDFIKILQMYKSPNASMEQLEHKLDTWRSV-LHNQVMPVLSIGPDASLY 267  
Qy 295 EAVYTLIKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLPRPSFLYRTIODLGIG 354  
Db 268 DAIKILHSRHLRPLVDPATGNVLYLTHKRLFLYINELPKPAYMQSLRELKIG 327  
Qy 355 TERDLAVVLETAPLITALDIFVDRRVSALPVVNECGOVVGLYSRFDVIHLAAQOTYNHLD 414  
Db 328 TNNIETADETTSITALKKFVRRVSALPLVDSGRVLDIYAKFDVINLAEKTYNDLD 387  
Qy 415 MSVGALRQRTLCLEGVLSQCPHESIGEVDIRAREQVHRLVLDVDTQHLGLVWSLSL 474  
Db 388 VSLRKANEHRNEPFGVKCNLDSELYTMEKIVRAEVHRLVVDENRKRVIIGIISLSL 447  
Qy 475 QALVLSAPG 483  
Db 448 LYLVLRPSG 456  
RESULT 4  
Q8SXT8  
ID Q8SXT8 PRELIMINARY; PRT; 906 AA.  
AC Q8SXT8;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE R622690P.  
GN SNF4AGAMMA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY084138; AAL89876.1; -;  
SQ SEQUENCE 906 AA; 99970 MW; C867D956F42D57F CRC64;  
Query Match 34.3%; Score 869.5; DB 5; Length 906;  
Best Local Similarity 53.1%; Pred. No. 9e-62;  
Matches 164; Conservative 73; Mismatches 71; Indels 1; Gaps 1;  
Qy 175 ELRKGAGIYMRFMQEHCTYDAMATSSKLVIFDTMLEIKKAFVALVANGVRAAPLWDSK 234  
Db 421 DLEEDDSQIFVKFFRHKCYDLIPTSAKLIVFDTQLLVKAFYALVYNGVRAAPLWDSK 480  
Qy 235 QSFVGMLTITDFILVLRHYRSPVQIYEIQHKTWREIYLGQCFKPLVSI SPNDSLF 294  
Db 481 QQFVGMLTITDFIKILQMYKSPNASMEQLEHKLDTWRSV-LHNQVMPVLSIGPDASLY 539  
Qy 295 EAVYTLIKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLPRPSFLYRTIODLGIG 354  
Db 540 DAIKILHSRHLRPLVDPATGNVLYLTHKRLFLYINELPKPAYMQSLRELKIG 599  
Qy 355 TERDLAVVLETAPLITALDIFVDRRVSALPVVNECGOVVGLYSRFDVIHLAAQOTYNHLD 414  
Db 600 TNNIETADETTSITALKKFVRRVSALPLVDSGRVLDIYAKFDVINLAEKTYNDLD 659  
Qy 415 MSVGALRQRTLCLEGVLSQCPHESIGEVDIRAREQVHRLVLDVDTQHLGLVWSLSL 474  
Db 660 VSLRKANEHRNEPFGVKCNLDSELYTMEKIVRAEVHRLVVDENRKRVIIGIISLSL 719

Qy 475 QALVLSAPG 483  
Db 720 LYLVLRPSG 728  
RESULT 5  
Q8SZS7  
ID Q8SZS7 PRELIMINARY; PRT; 1400 AA.  
AC Q8SZS7;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE LD22662P.  
GN SNF4AGAMMA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY070541; AAL48012.1; -;  
SQ SEQUENCE 1400 AA; 152380 MW; 411B93CG6B9EC7AF CRC64;  
Query Match 34.3%; Score 869.5; DB 5; Length 1400;  
Best Local Similarity 53.1%; Pred. No. 9e-62;  
Matches 164; Conservative 73; Mismatches 71; Indels 1; Gaps 1;  
Qy 175 ELRKGAGIYMRFMQEHCTYDAMATSSKLVIFDTMLEIKKAFVALVANGVRAAPLWDSK 234  
Db 915 DLEEDDSQIFVKFFRHKCYDLIPTSAKLIVFDTQLLVKAFYALVYNGVRAAPLWDSK 974  
Qy 235 QSFVGMLTITDFILVLRHYRSPVQIYEIQHKTWREIYLGQCFKPLVSI SPNDSLF 294  
Db 975 QQFVGMLTITDFIKILQMYKSPNASMEQLEHKLDTWRSV-LHNQVMPVLSIGPDASLY 1033  
Qy 295 EAVYTLIKNRHRLPVLDPVSGNVHLTHKRLKFLHIFGSLPRPSFLYRTIODLGIG 354  
Db 1034 DAIKILHSRHLRPLVDPATGNVLYLTHKRLFLYINELPKPAYMQSLRELKIG 1093  
Qy 355 TERDLAVVLETAPLITALDIFVDRRVSALPVVNECGOVVGLYSRFDVIHLAAQOTYNHLD 414  
Db 1094 TNNIETADETTSITALKKFVRRVSALPLVDSGRVLDIYAKFDVINLAEKTYNDLD 1153  
Qy 415 MSVGALRQRTLCLEGVLSQCPHESIGEVDIRAREQVHRLVLDVDTQHLGLVWSLSL 474  
Db 1154 VSLRKANEHRNEPFGVKCNLDSELYTMEKIVRAEVHRLVVDENRKRVIIGIISLSL 1213  
Qy 475 QALVLSAPG 483  
Db 1214 LYLVLRPSG 1222  
RESULT 6  
Q9BHL6  
ID Q9BHL6 PRELIMINARY; PRT; 372 AA.  
AC Q9BHL6;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Y11B2A.8 protein.  
GN Y11B2A.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RA Sulston J.E.;
RP SEQUENCE FROM N.A.
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL132904; CAC35836.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 372 AA; 41376 MW; 81A39670877167DF CRC64;

Query Match 29.0%; Score 736.5; DB 5; Length 372;
Best Local Similarity 49.8%; Pred. No. 1e-51;
Matches 148; Conservative 63; Mismatches 83; Indels 3; Gaps 2;

QY 188 MQEHTCDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTITDFI 247
Db 1 MKAHKCYDLPTSSKLVFDTMLPVKFAFYALVINGVRAAPLWDTNQRFTGMLTITDFI 60
QY 248 LVLRHYV--RSPVLQVIEIQEHKIEWTREIY-LQGCFKPLVSISSPNDLSFEAVYTLIKNR 304
Db 61 KILCKHYKDNSEIRALEDDQIISHWRDQFELDTLRFVYIDNESLHRAVELLCESK 120
QY 305 IHRPLVLPVSGNVLIHTHRLKLFHIFGSLLPSPFLYRTIQDGLGIGTFRLAVYLE 364
Db 121 VHRPLVLDKRTGNTITVILTHKRMKFLSLYMRDLPRSEMSCTPRELGIGAWGDILCOHV 180
QY 365 TAPILTALDIFDVRVSALPVVNECGQVVGLYSRFDVIHAAQTYNHLDSVGEALRQR 424
Db 181 DTPITHDALEFLKRNRYALPLFDENGVRVYIAKFQVDSLAAESYDKLDCTVOEALQHR 240
QY 425 TLCLEGLVSCOPHESGEVIDRIAREQVHRLVLVDETOHLGLVWSLSDILQALVSP 481
Db 241 SENFEGVQTCLETDSLQFVLEAIVKAEVHRLVITDQKKVGVWSLSDILKALVLDP 297

RESULT 7
Q8SSV7 PRELIMINARY; PRT; 577 AA.
AC Q8SSV7
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SNF4/AMP-activated protein kinase gamma subunit.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lehmann R., Baugart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115599; AAL92329.1; -.
SQ SEQUENCE 577 AA; 64209 MW; 3FBC56EA649B25A9 CRC64;

Query Match 22.0%; Score 559.5; DB 5; Length 577;
Best Local Similarity 31.3%; Pred. No. 4.7e-37;
Matches 151; Conservative 100; Mismatches 178; Indels 53; Gaps 17;

QY 33 NSSSWPSPATYSSSIRKRAKALRWTRQKSVGEPPGQGEPP-----RSRPAESTG 88
Db 105 NSIQPSPFISSQ--DGLVTVDPVLDVADGKGNKESQSPPPGNDQNLNNMFPKDKITS 162

```

```

QY 89 LEATFPKTT-----PLAQ--ADPA-GVCTPPTGWDCLPSDCTASAAGSSDDVE 134
Db 163 LPSTDNKSNTNNNNENPLKQTISSPFSKSTTTTSTSTTTTSPSSUSSNNNNNSN 222
QY 135 LATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDLELRK---PGAQIYRFRQEH 191
Db 223 SNNYINHSSISTVSGLEN-----LNLKSGIKKI--DSETEKYIEEGKQVFNFLKGH 274
QY 192 TCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMLTITDFILVLH 251
Db 275 TCYDVIPISGKVVVLDTKLAVSAFYALEENGISAKPLWNSEQHDTGMITVSDIFIDLL 334
QY 252 RYRSPLVQ--IYEIQHKIET-WREIYLOGCFKP--LVSISPNDSLPEAVYTLIKNRH 306
Db 335 YYYKKPKSNFIQDMGIHRIETFWREISVE---RPSSLISTEPEPINLDAASLLCCKIHI 391
QY 307 RLPVLDPVSGN-VLIHTLHRLKLFHIFGSLLPSPFLYRTIQDGLGIGTFRLAVYLET 365
Db 392 RLPVVDKKTNSILHILTHSRILAFPMKSPQLPE-KLLSIPIGSLGIGTFATVTVVWTH 450
QY 366 APILTALDIFDVRVSALPVV-NECGQVVGLYSRFDVIHAAQTYN-----HLDMS 416
Db 451 TPLVEVLELSEKKISAVPIIDSETSKIVDVYSKSDVTLMKSGKILSPSDNLNLPVHQVLS 510
QY 417 VGEALRQRTLCLEGLVSCOPHESGEVIDRIAREQVHRLVLVDETOHLGLVWSLSDILQA 476
Db 511 TPTKLWQRP--EQIVTCTRFKGLGVIERCIRKVRHRLVCIIDSSKKVEGILSLDILNY 567
QY 477 LV 478
Db 568 LL 569

RESULT 8
002168 PRELIMINARY; PRT; 423 AA.
ID 002168
AC 002168;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 47.5 Kda protein.
GN T20F7.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97550; AAK18981.2; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Hypothetical protein.
SQ SEQUENCE 423 AA; 47451 MW; 6997065D515E7B21 CRC64;

```



[illegible]



```
QY 216 FFALVANGYRAAPLWDSKKQSFVGMT-----ITDFILVL--HRYRSPLVQIYEIEQHKI 269
Db 90 VHALSQHGHIAAVVTNTDKQAECVFMGMHCLTALLVAAGNREVASKTL-----V 140
QY 270 ETWREIYLOGCFKPLVSPNDSEFAVYTLIKNRIHRLPVLDPVSGNVHLHILTHKR 326
Db 141 EFLKEIGSQN----TICSGVQNSVWEAANIISHNKISFVPIFDTIIPKPGTPLYFLTPRM 196
QY 327 LLK-----FLHIFGSLPRPSEFLYRTIQDLGIGTERD--LAVVLETAPILTA 371
Db 197 ILQETVLKLSDFGDAILLHV-----RQATLDQKKIGTWDDVLKICLNTT-IEEA 245
QY 372 LDIFVDRRVSALPVVNECCQVVGLYSRFDVI-HLAAQOTYNHLDMSVGEALRQRTLCLEG 430
Db 246 IKLMSERKXSTIPVVVNDFKQIVNMLARKDILILEINSHQCGNFHDK-----LKEPVKILQS 300
QY 431 VLS-----COPHESLGEVIDRIAREQVHRLVLYDETOHLLGVVSLSDIL 474
Db 301 LQSRLYGRSSYTFETVAKMNTSDKSSLPIDEGRILAVVSCDIL 348
```

Search completed: June 6, 2003, 11:02:49  
Job time : 66.373 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 29.1173 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2338  
Sequence: 1 MEPCLEHALRTPSSWISLG.....LSDILQALVLPAGIDALGA 489

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1080.5	42.6	328	2	US-08-878-989-7
2	1080.5	42.6	328	4	US-09-272-796-7
3	1028	40.5	330	4	US-09-359-161-6
4	1017	40.1	331	2	US-08-878-989-21
5	1017	40.1	331	3	US-09-101-146-64
6	1017	40.1	331	4	US-09-272-796-21
7	457	18.0	322	4	US-09-359-161-7
8	214	8.4	379	4	US-09-359-161-5
9	190	7.5	373	4	US-09-359-161-3
10	102	4.0	2273	4	US-09-426-998-5
11	101	4.0	187	4	US-09-199-637A-287
12	97	3.8	852	2	US-09-070-060-3
13	97	3.8	852	3	US-09-357-746-3
14	96	3.8	830	4	US-09-562-737-33
15	94.5	3.7	1285	1	US-07-582-945-2
16	94.5	3.7	1285	2	US-08-453-141-2
17	94.5	3.7	1285	3	US-08-293-314-2
18	94	3.7	830	4	US-09-562-737-37
19	94	3.7	2353	4	US-08-984-709A-50
20	93.5	3.7	1208	4	US-09-463-702A-2
21	91.5	3.6	854	2	US-09-070-060-4
22	91.5	3.6	854	3	US-09-357-746-4
23	91.5	3.6	2972	4	US-09-579-181-2
24	91.5	3.6	3118	4	US-09-579-181-1
25	90.5	3.6	443	4	US-09-134-001C-3767
26	90.5	3.6	599	2	US-08-426-125-1
27	90.5	3.6	599	2	US-08-455-355-1

Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 12, Appli  
Sequence 12, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-878-989-7  
; Sequence 7, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PENITUT01

; CLONE: 1452972  
US-08-878-989-7

Query Match 42.6%; Score 1080.5; DB 2; Length 328;  
Best Local Similarity 63.3%; Pred. No. 2.3e-111;  
Matches 200; Conservative 67; Mismatches 48; Indels 1; Gaps 1;

QY 169 KLGWDF-LRKPCAQIYRMFQHEQTOYDAMATSKLVIPTMTLETKKAFFALVANGVRAA 227  
DB 4 KLEFEDAEVDESSEGVYMFMRSHKCYDIPTSSKLWVFDTTLQVKKAFFALVANGVRAA 63  
QY 228 PLWDSKQSFVGMILTITDFILVHLHYRSPVQIYEIOHQKIETWREIYLQGCFFKPLVSI 287  
DB 64 PLWESKQSFVGMILTITDFILVHLHYRSPVQIYELEBHKIETWREIYLQETFKPLVNI 123  
QY 288 SPNDSLEAVYTLIKNRHRLPVLDPVSGVNLHILTHKRLKFLFGLSLLPRPSFLYRT 347  
DB 124 SPNDSLEAVYTLIKNRHRLPVLDPVSGVNLHILTHKRLKFLFGLSLLPRPSFLYRT 347  
QY 348 IQDLGIGTFRDLAVLETAITLTDIFVDRRVSALPVVNEGVGVGLYSRFDVHILAAQ 407  
DB 184 LDELGIGTGHNTAFTHPTPIIKALNIFVRRISALPVVDESQKVDIYSKFDVINLAAE 243  
QY 408 QTYNHLDMVSGEALQRITCLLEGVLSQPHESLGEVIDRIAREQVHRLVLVDQTHLLGV 467  
DB 244 KTYNNLDITVTQALQHRSOYFEGVYKCNKLEILETTIVDRIVRAEVHRLVVVNEADSI 303  
QY 468 VLSLSDILQALILTPAG 483  
DB 304 ISLSLILQALILTPAG 319

RESULT 2

US-09-272-796-7  
Sequence 7, Application US/09272796  
Patent No. 6207148

GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl G.  
APPLICANT: Lal, Preeti  
APPLICANT: Goli, Surya K.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/272,796  
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: 08/878,989  
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0321 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PENITUT01  
; CLONE: 1452972  
US-09-272-796-7

Query Match 42.6%; Score 1080.5; DB 4; Length 328;  
Best Local Similarity 63.3%; Pred. No. 2.3e-111;  
Matches 200; Conservative 67; Mismatches 48; Indels 1; Gaps 1;

QY 169 KLGWDF-LRKPCAQIYRMFQHEQTCYDAMATSKLVIPTMTLETKKAFFALVANGVRAA 227  
DB 4 KLEFEDAEVDESSEGVYMFMRSHKCYDIPTSSKLWVFDTTLQVKKAFFALVANGVRAA 63  
QY 228 PLWDSKQSFVGMILTITDFILVHLHYRSPVQIYEIOHQKIETWREIYLQGCFFKPLVSI 287  
DB 64 PLWESKQSFVGMILTITDFILVHLHYRSPVQIYELEBHKIETWREIYLQETFKPLVNI 123  
QY 288 SPNDSLEAVYTLIKNRHRLPVLDPVSGVNLHILTHKRLKFLFGLSLLPRPSFLYRT 347  
DB 124 SPNDSLEAVYTLIKNRHRLPVLDPVSGVNLHILTHKRLKFLFGLSLLPRPSFLYRT 347  
QY 348 IQDLGIGTFRDLAVLETAITLTDIFVDRRVSALPVVNEGVGVGLYSRFDVHILAAQ 407  
DB 184 LDELGIGTGHNTAFTHPTPIIKALNIFVRRISALPVVDESQKVDIYSKFDVINLAAE 243  
QY 408 QTYNHLDMVSGEALQRITCLLEGVLSQPHESLGEVIDRIAREQVHRLVLVDQTHLLGV 467  
DB 244 KTYNNLDITVTQALQHRSOYFEGVYKCNKLEILETTIVDRIVRAEVHRLVVVNEADSI 303  
QY 468 VLSLSDILQALILTPAG 483  
DB 304 ISLSLILQALILTPAG 319

RESULT 3

US-09-359-161-6  
Sequence 6, Application US/09359161A  
Patent No. 6342656

GENERAL INFORMATION:

APPLICANT: Bradford, Kent J.  
APPLICANT: Dahal, Peetambar  
APPLICANT: Yang, Hong  
APPLICANT: Cooley, Michael  
APPLICANT: Downie, Bruce  
APPLICANT: Gee, Oliver  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses  
FILE REFERENCE: 023070-095900US  
CURRENT APPLICATION NUMBER: US/09/359,161A  
CURRENT FILING DATE: 1999-07-21  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn ver. 2.1  
SEQ ID NO 6  
LENGTH: 330  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: gamma subunit of AMP-activated protein kinase  
OTHER INFORMATION: (AMPK-gamma)

Query Match 40.5%; Score 1028; DB 4; Length 330;

Best Local Similarity 62.3%; Pred. No. 1.6e-105;  
Matches 197; Conservative 56; Mismatches 63; Indels 0; Gaps 0;

```
QY 165 APPFKLGHDDDELKPGAGIYWRPQHEHCYDAMATSSKLVIFDTMLEIKKAFALVANGV 224
Db 9 APAPENESQTPSPNSSVYTTMKSHRCYDLIPTSSKLVVFDTSLOVKKAFALVTNGV 68
QY 225 RAAPLWDSKKQSFVGMITITDFILVLRHYRSPVQIYEIEQHKTETWREIYLOGCFKPL 284
Db 69 RAAPLWDSKKQSFVGMITITDFILVLRHYRSPVQIYEIEQHKTETWREIYLOGCFKPL 128
QY 285 VSISPNDLSFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLKELHIFGSLPRPSFL 344
Db 129 VCISPNASLFDVAVSSLRNRIHRLPVLDPVSGNVLHILTHKRLKELHIFGSLPRPSFL 188
QY 345 YRIODLIGTFRDLAVVLEFAPILTALDIFVDRVRSALPVNCGQVGLYSRFDVHL 404
Db 189 SKSLEELQIGYIANIAWRTTTPVVALGIFVQHRVSALPVNCGQVGLYSRFDVHL 248
QY 405 AAQOTYNHLDMSVGEALRQRTLCLEGLVSCQPHESLGEVIDRIAREQVHRLVLDVETQHL 464
Db 249 AAEKTYNNLDVSVTKALQHRSHYFEGVLKCYLHETLEALINRLVEAEVHRLVVDVDEHDV 308
QY 465 LGVWLSLILQALVLS 480
Db 309 KGIVLSLILQALVLT 324

RESULT 4
US-08-878-989-21
; Sequence 21, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1335856
US-08-878-989-21

Query Match 40.18; Score 1017; DB 2; Length 331;
Best Local Similarity 64.8%; Pred. No. 2.7e-104;
Matches 193; Conservative 53; Mismatches 52; Indels 0; Gaps 0;

QY 183 IYRMFWQHEHCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPLWDSKKQSFVGMILT 242
Db 28 IYSPMKSHRCYDLIPTSSKLVVFDTSLOVKKAFALVTNGVRAAPLWDSKKQSFVGMILT 87
QY 243 IYDFILVLRHYRSPVQIYEIEQHKTETWREIYLOGCFKPLYSISPNDLSFEAVYTLIK 302
Db 88 ITDFINILHRYKSAVQIYEIEHKTETWREIYLOGCFKPLVCISPNASLFDVAVSSLR 147
QY 303 NRHRLPVLDPVSGNVLHILTHKRLKELHIFGSLPRPSFLYRTODLIGTFRDLAVV 362
Db 148 NKIHLRPVLDPESGNTLYLTHKRLKELHIFGSLPRPSFLYRTODLIGTFRDLAVV 207
QY 363 LETAPILTALDIFVDRVRSALPVNCGQVGLYSRFDVHIAAQOTYNHLDMSVGEALR 422
Db 208 RTTTPVVALGIFVQHRVSALPVNCGQVGLYSRFDVINLAEXTYNNLDVSVTKALQ 267
QY 423 QRTLCLEGLVSCQPHESLGEVIDRIAREQVHRLVLDVETQHLQVWLSLILQALVLS 480
Db 268 HRSHYFEGVLKCYLHETLETLINRLVEAEVHRLVVDVDEHDVKGIVLSLILQALVLT 325

RESULT 5
US-09-101-146-64
; Sequence 64, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-101-146-64

Query Match 40.18; Score 1017; DB 3; Length 331;
```

```

Best Local Similarity    64.8%;   Pred. No. 2.7e-104;
Matches 193; Conservative 53; Mismatches 52; Indels      0; Gaps      0;

QY          183 IYNRFMQEHTCDAMATSSKLVIFDTMLEIKKAFFALVANGVRAPLWDSKKQSFGMLT 242
              :|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
Db          28 VYTSEPMKSHRCYDLPTTSKSLVDFDSLQVKKAFAFALTNGVRAAPLWDSSKKSGFYGLT 87

QY          243 ITDFILVLHRYRSPLVOIYEIOHKITETWREIYLGGCFKLVSISPNDLSLFNAVTLIK 302
              ||||::||||::||||::||||::||||::||||::||||::||||::||||::||||
Db          88 ITDFINILHRYKKSALVOIQVELEBHKITETWREVIQLDSFKPLVCISPNASLFDVAVSSLIR 147

QY          303 NRIHRLPVLDPSGVNVLHLTHKKLLKFLHFCSSLPRPSPLYRTIQDGLIGTFRDLAVV 362
              |:|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db          148 NKTRHLPVDPESGNTLVLTTHKKLKFLKLFTEPKPEPFMSLSLEELQIGTYANIAYV 207

QY          363 LETAPTILTALDITFDVRRYSALPVVNQCQVVGLYSRFDVHLLAQOITYNHLMDSVGEALR 422
              :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db          208 RTTPVYVALGIFGVQRHSALPVVDKGRVDDIYSKFVINLAKEKTYNNLDSVTTKALQ 267

QY          423 QRTLCEGLVSCQPHESELGEVIDRIAREQVHRVLVLVDETOHLLGWLSLSLILQAALVLS 480
              |:||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db          268 HRSHYEGVLCYLKHETLETIIINSLVAEVRHLVVVDENDVVGIVSLSLILQAALVLT 325


RESULT 6
US-09-272-796-21
: Sequence 21, Application US/09272796
: Patent No. 6207148
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corley, Neil C.
: APPLICANT: Guegler, Karl G.
: APPLICANT: Lal, Preeti
: APPLICANT: Goli, Surya K.
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
: TITLE OF INVENTION: KINASES
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/272,796
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/878,989
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J J
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0321 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 331 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:

```

```
; LIBRARY: GenBank
; CLONE: 1335856
US-09-272-796-21

Query Match          40.1%; Score 1017; DB 4; Length 331;
Best Local Similarity 64.8%; Pred. No. 2.7e-104;
Matches 193; Conservative 53; Mismatches 52; Indels 0; Gaps 0;

QY      183 IYRFRMOEHCYDAMATSKLIVFDNMLEIKKAFALVANGVRAAPLWDSKKOSFVGMLT 242
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      28 VYTSMFKSHRCVDLIPTSSKLVPEDTSLOVKKAFFALTNGVRAAPLWDSKKQSFGVGLT 87

QY      243 ITDFILVLHRYPSPLVOIYEIOHKTIETWREIYLQGCFKP LVSISPNDLSFEAVNTLIK 302
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      88 ITDPINILHYKYKSALVOIYELEBKTIETWREIYQLDSFKPLVCISPNASLDFDVSSLIR 147

QY      303 NRIHRPVLDPVSGNWHLTHKRLLKFHFICGSLLRPRSPLYRTIQDLIGITGFRLDAVV 362
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      148 NKIHRLPVDPESGNTLYLTHKRLLKFLKLFTEFPKPEMSKSLEEQLIGTYANIAMV 207

QY      363 LETAPILTALDIYDRRVSA LPVYNCGGVGLYSRFDYVIHLAAOQTYNHLMDSVGEALR 422
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      208 RTTPPVVALGIFVGHRSVPVVDCKRGVVDIYSKFDVINLAAEKTYNNLDVSVTKALQ 267

QY      423 QRTLCEGLVSCQPESHSEVIDRIAEQVHRVLVLVDQTQHLLGWLSLDILOALVLS 480
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      268 HRSHYFGLVKCYLKHETLETINLVAEVRHVLUVVVDENVKGIVSLSLDILOALVLT 325


RESULT 7
US-09-359-161-7
; Sequence 7, Application US/09359161A
; Patent No. 6342856
; GENERAL INFORMATION:
; APPLICANT: Bradford, Kent J.
; APPLICANT: Dahal, Peetambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; TITLE OF INVENTION: To Stress Conditions in Plants
; FILE REFERENCE: 023070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: yeast sucrose nonfermenting protein kinase 1
; OTHER INFORMATION: kinase subunit (SNF1)
US-09-359-161-7

Query Match          18.0%; Score 457; DB 4; Length 322;
Best Local Similarity 34.2%; Pred. No. 4.7e-42;
Matches 106; Conservative 66; Mismatches 108; Indels 30; Gaps 6;

QY      186 RFMOEHCTCDAMATSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKOSFVGMLTTID 245
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      24 KFLNSKSTSYDLVPVSRKLIVLDTSLWKKSLANVLQNISVSA PLWDSKISRAGLLTTID 83

QY      246 FILVLHRYPSPLVOIYEIOHKTIETWREIYLQGCFKP L-----VSISPNDLSLF 294
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      84 FINVIQYFSNP-----DKELVDKLQLDG-LKDIERALGVQDOLDTASHIPSRPLF 133

QY      295 EAYVTLIKNIHPLPVLPDPVGN----VULHTHKLKLFHIEGSLLRPRS--FLYRTIQ 349
       || :||:||||:| |:| | | | | | | | | | | | | | | | | | | | | | | |
Db      134 EACLKMLESSGRPIHQDEBETHREVSVTLQVRIKLKV---ALNCRETHFLKIPIG 189
```

	Query Match	18.0%;	Score 457;	DB 4;	Length 322;
	Best Local Similarity	34.2%;	Pred. No. 4.7e-42;		
	Matches 106;	Conservative 66;	Mismatches 108;	Indels 30;	Gaps 6;
Qy	186	RMOEHTCYDAMATSKLVIFDITMLEIKTKAFFALVANGVRAAPLWDSKKQSFVGMLTITD	245		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
Db	24	KFLNSKTSYDVLVPVSKVLIVDTLSLLVKKSLNVLQNSIVSAPLWDSKTSRRAGLLTITD	83		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
Qy	246	FILVLHRYRSPVLQVIEYEQHKETWREIYLGQCFKPL-----VTSIPNDSLF	294		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
Db	84	FINVIQYFSNP-----DKELVDKLQLDG-LKOIERALGVQDQDPAISHPSRPLF	133		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
Qy	295	EAVYTLKNIHRLPVLDPVSGN-----VLHILTHRLKRLKFLHIEGSLLLPRS-FLYRTIQ	349		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			
Db	134	EACLKMLLESSGRIPILQDSETHREIVSVLTQVRLKVF---ALNCSETHFLKIDIP	189		
		: :     :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :			

```

Qy 350 DLGIGTRDLAVVETAPILFALDI FVDRKVSALPVVNECGGVGLYSRFDVTHLAAQQT 409
    ||| : : || : : ||| : : || : : ||| : :
Db 190 DLNIITODNMKSCQMTTPVIDIOMLTQGRVSSVPIIDENGYLINVVEAYDVLGLKGGI 249
    ||| : : ||| : : ||| : : ||| : : ||| : :
Qy 410 YNHLDMSVGEALRQRTLCLGVLSCQPSHESGEVIDRIAREQVHRLVIVDETQHLLGVVS 469
    ||| : : ||| : : ||| : : ||| : : ||| : :
Db 250 YNDLSLVGEALMRKRSDDFEGVYTCTKNDKLTIMDNIRKRVHRFFVYDDVGRLGVLT 309
    ||| : : ||| : : ||| : : ||| : : ||| : :
Qy 470 LSDILQALV 479
    |||| : ||
Db 310 LSDILKYILL 319
    |||| : ||

RESULT 8
US-09-359-161-5
; Sequence 5, Application US/09359161A
; Patent No. 6342656
; GENERAL INFORMATION:
; APPLICANT: Bradford, Kent J.
; APPLICANT: Dahal, Peetambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; TITLE OF INVENTION: to Stress Conditions in Plants
; FILE REFERENCE: 023070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
; FEATURE:
; OTHER INFORMATION: Phaseolus vulgaris L. Pv42
US-09-359-161-5

```

```
; FILE REFERENCE: ORT-1057
; CURRENT APPLICATION NUMBER: US/09/426,998
; CURRENT FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PATENTIN VER. 2.0
; SEQ ID NO 5
; LENGTH: 2273
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-426-998-5

Query Match      4.0%; Score 102; DB 4; Length 2273;
Best Local Similarity 19.5%; Pred. No. 0.4;
Matches 122; Conservative 77; Mismatches 194; Indels 234; Gaps 30;

QY  2  EPGLLEALRRTPSSLSGGSEHQEMSFLEQENSSWPSPAVTSSEIRIKRRAKALRWT 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1076 EPGANIEMKSPSARS-----SPSPWS-----AASWTS-----RSSRNSLG 1114

QY  62  RQKSVEEGEPGQ-----GEGPRSPAAESTGLEATPKTTPPLAQADPAGVTPPTG-- 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 RAPSLKRSPSGERRSLLSGEQSQSQDOESSEE-----RASPAGSDHRRGSL 1164

QY  114  -----WDCLPSDC-----TASAGSSTDVE-----LATEFPATEAME 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 EREAKSFD-LPDTLQVPLHRTASGRGSAHEQDNGKSASGRALARALRPDDPLDGD 1223

QY  147  CELEGLIE--ER-----PALCLSPQAPFPKLGWDELKPKGA--QIYMRPMQHTC 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1224 ADDEGNLSKGERVARIKAPACLERDS-----WSAYIFPPQSRFLLCHRIITHKM 1277

QY  194  YDAMATSSKLVIFDTMLFI-----KKAFFALVANGVRAAPL-----W 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1278 FDHVVL---VIFLNCITIAMERKPIDPSAERIFLTLNFIYTAFLAEMTVKVVALGW 1334

QY  231  DSKKQSV-GMLTITDFILVIRYKSHVQIYEIEQHKIETWRIYIYQCGCFPLYSISP 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 CFGEQAYLRSWNLVGLVILVSDILVSNWSUSQTKILGMLRVLRLLRTLRPLRVIS 1394

QY  290  NDSIFEAVYTLIKNRIHRLPVLDPVSGNVHLIHLTHKRLKFLHIFGSL---LPPSP-L-- 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1395 AQGLKLVETLMS-----LKPI-GNIVVI-----CCAFFIIFGILGVLFKGFVC 1441

QY  345  -----YRTIQDIGTGTFRDLAVVLETAPIALTDALFIDVRRVYSALPVV 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 QGEDTRNLTKNSDCAFASYRWRIK--YNFDNIQAQIMSLFVLASKDGMFI----- 1491

QY  387  NECQVGLYSRFDVTHILAAQYTNH-----IDMSVGEALRQKTLIC 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1492 -----MYDGLDVGVDQDQIMNINIWMILYFISPLIIVAFFVILNMEVG-----VV 1546

QY  428  LEGVLSOPHESLGEV-----IDRIARQ-----VHRL 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1537 VENFHKCRQHEEERREKRUPKLEKKRRSEKQMAEACKPYYSIDYSRFLLVHRL 1596

QY  456  VLVDETQHLI-----GVVSLSDIIOAL 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1597 C-----TSHYLDLFIITGVIGLNVVTMAM 1619

RESULT 11
US-09-199-637A-287
; Sequence 287, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Kaimo, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsoukalas, John
```

```
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-287

Query Match      4.0%; Score 101; DB 4; Length 187;
Best Local Similarity 25.0%; Pred. No. 0.0069;
Matches 44; Conservative 14; Mismatches 46; Indels 72; Gaps 10;

QY  38  PSPAVTSS-----SERIRKRRAKALRWTQ-----KSVEEGEPGQG 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19  PAPCTTRSCPPRSRGRTSRAGSDR--RGRRANGARRWTRRLPPRGRSLADAAPCA 77

QY  76  EGPSRPAEST-GLEAT-----FPKTTPLAQAD-----PAGVGTPTPTGWD---C 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  78  AASRARPRASSTPGYRSTWKPLRRFPFRSPCCATGSRSCRACRGSACVPGGWSGTGC 137

QY  117  LPSPD-----TASAGSSTDDELATEFPATEAMECEGLEELERPALCLSPQAFP 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  138  APCSRGRPCAAP-----PPSPAW-----PARSSAGSAPSP 167

RESULT 12
US-09-070-060-3
; Sequence 3, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Hustad, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 852 amino acids
; TYPE: amino acid
```

; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-070-060-3

Query Match 3.8%; Score 97; DB 2; Length 852;  
Best Local Similarity 22.1%; Pred. No. 0.26;  
Matches 64; Conservative 40; Mismatches 114; Indels 72; Gaps 15;

QY 29 LEQENSSWSPAVTSSSIRGKRRAKALRWTRQKSVGEPPGCGEGPR----SKPAA 84  
DB 134 LESEVNTGTTCSSESQNDGSRKDETRVSTNGS--DQPEDAGAGENRRVSGNNSPSL 192  
QY 85 ESTGLRATPP----KTTPLAQADPAGVGTPTGWCDCPLSDCTASAAGS-----STDDVE 134  
DB 193 SNGGPKSPRPSPRPPTPRPASVNGSPSA----TSESDGSSTGSLPTNTNTNSE 248  
QY 135 LATE---FPATAECELEGLLEERPALCLSP--QAPPPKLGWDELRKPGAOIYRPMQ 189  
DB 249 GATSGLIILPT-----ISGSGGRP---LNPVTQAPLPP--GWEQRVDQGRVYVDHVE 298  
QY 190 EHTCYDAMATSKLVIFDTMLEIKKAFVALVANGVRAAPL---WDSKKQSFVGMILTIDF 246  
DB 299 KRTTWD-----RPEPLPGWERRVDN--MGRIVYVDH 328  
QY 247 ILVLHRYRSPLVQIYEIQHKIETWREIYLQCFKPLVS---ISPNDSLF 294  
DB 329 FTRTTWQRPITLESVRNYEQWLQ---RSQLOGAMQQFNQRFIYGNQDLF 375

RESULT 13

US-09-357-746-3  
; Sequence 3, Application US/09357746  
; Patent No. 6087122  
; GENERAL INFORMATION:  
; APPLICANT: ZENECA Limited  
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE  
; FILE REFERENCE: PHM.70312.N1  
; CURRENT APPLICATION NUMBER: US/09/357,746  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: US No. 608712209/070.060  
; EARLIER FILING DATE: 1998-04-30  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 852  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-357-746-3

Query Match 3.8%; Score 97; DB 3; Length 852;  
Best Local Similarity 22.1%; Pred. No. 0.26;  
Matches 64; Conservative 40; Mismatches 114; Indels 72; Gaps 15;

QY 29 LEQENSSWSPAVTSSSIRGKRRAKALRWTRQKSVGEPPGCGEGPR----SKPAA 84  
DB 134 LESEVNTGTTCSSESQNDGSRKDETRVSTNGS--DQPEDAGAGENRRVSGNNSPSL 192  
QY 85 ESTGLEATPP----KTTPLAQADPAGVGTPTGWCDCPLSDCTASAAGS-----STDDVE 134  
DB 193 SNGGFKSPRPSPRPPTPRPASVNGSPSA----TSESDGSSTGSLPTNTNTNSE 248  
QY 135 LATE---FPATAECELEGLLEERPALCLSP--QAPPPKLGWDELRKPGAOIYRPMQ 189  
DB 249 GATSGLIILPT-----ISGSGGRP---LNPVTQAPLPP--GWEQRVDQGRVYVDHVE 298  
QY 190 EHTCYDAMATSKLVIFDTMLEIKKAFVALVANGVRAAPL---WDSKKQSFVGMILTIDF 246  
DB 299 KRTTWD-----RPEPLPGWERRVDN--MGRIVYVDH 328  
QY 247 ILVLHRYRSPLVQIYEIQHKIETWREIYLQCFKPLVS---ISPNDSLF 294

DB 329 FTRTTWQRPITLESVRNYEQWLQ---RSQLOGAMQQFNQRFIYGNQDLF 375

RESULT 14

US-09-562-737-33  
; Sequence 33, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDU Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 830  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-33

Query Match 3.8%; Score 96; DB 4; Length 830;  
Best Local Similarity 24.5%; Pred. No. 0.32;  
Matches 46; Conservative 20; Mismatches 74; Indels 48; Gaps 9;

QY 2 EPGLEHALRRTPSWSLGSGSEHQMSFLEQENSSWSPAVTSSSIRGKRRAKALRW 61  
DB 228 DPGTEHDLRSHSGGIEGRSSQELK-----SPG--SDSEDALGARLGRMISMI 274  
QY 62 RQKSVGEPPGCGEGPRSPRAAESTGLEATFKTTPLAQADPAGVGTPTGWCDCPLSDC 121  
DB 275 SETELELNDGGSSSG--RSQHLTNSIB--EARSAPSEPESEP--LHEPPRTATFLP--- 327  
QY 122 TASAAGSSTDDVELATEFPATEAWECEL-----EGLLEER-----PALCL 161  
DB 328 -----VGQDDVN--SEYESSEWEPDLSADADYPHLLSNLYSMISEGSPIDCFQOCL 379  
QY 162 SPQAPPK 169  
DB 380 SPAERLPE 387

RESULT 15

US-07-582-945-2  
; Sequence 2, Application US/07582945  
; Patent No. 5369019  
; GENERAL INFORMATION:  
; APPLICANT: TAKKER FOGED, Niels  
; APPLICANT: PETERSON, Svend  
; TITLE OF INVENTION: A PASTEURELLA VACCINE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/582,945  
; FILING DATE: 19901012  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 11:02:59 ; Search time 40.2504 Seconds  
(without alignments)  
1254.259 Million cell updates/sec

Title: US-09-826-581-6  
Perfect score: 2538  
Sequence: 1 MEPGLEHALRRTPSWSSGG.....LSDTLQALVLSAGIDALCA 489

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US50\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2538	100.0	489	10	US-09-826-581-6
2	1017	40.1	344	10	US-09-925-297-461
3	892.5	35.2	1207	9	US-10-108-605-71
4	115	4.5	980	9	US-10-108-605-195
5	114	4.5	91	10	US-09-864-761-33979
6	105.5	4.2	1134	9	US-10-001-873-50
7	101	4.0	187	9	US-09-975-719-287
8	100.5	4.0	440	9	US-09-738-626-6006
9	100	3.9	1276	10	US-09-272-809-2
10	99.5	3.9	701	9	US-10-052-092-12
11	99.5	3.9	701	9	US-10-211-613-1
12	97.5	3.8	1139	9	US-09-971-490-5
13	97	3.8	739	9	US-10-097-534-10
14	96	3.8	830	9	US-10-211-962-33
15	96	3.8	1604	10	US-09-746-491-8
16	95.5	3.8	622	9	US-09-738-626-4919
17	94	3.7	830	9	US-10-211-962-37
18	93.5	3.7	419	9	US-10-257-963-12
19	93.5	3.7	419	10	US-09-175-254-3

20	93.5	3.7	2462	9	US-09-819-104A-5	Sequence 5, Appli
21	92.5	3.6	529	12	US-10-042-417-44	Sequence 44, Appl
22	92.5	3.6	1069	9	US-10-007-270-24	Sequence 24, Appl
23	92	3.6	595	9	US-10-052-092-31	Sequence 31, Appl
24	92	3.6	758	10	US-09-904-987-5	Sequence 5, Appli
25	91.5	3.6	660	10	US-09-853-033-8	Sequence 8, Appli
26	91	3.6	660	10	US-09-853-033-4	Sequence 4, Appli
27	90.5	3.6	909	9	US-10-226-296-4	Sequence 4, Appli
28	90.5	3.6	1169	10	US-09-815-242-13693	Sequence 13693, A
29	90.5	3.6	1618	9	US-09-963-875-1	Sequence 1, Appli
30	90.5	3.6	1618	9	US-10-136-891-2	Sequence 2, Appli
31	90.5	3.6	1618	9	US-10-120-687-1	Sequence 1, Appli
32	90.5	3.6	2016	9	US-09-896-994-2	Sequence 2, Appli
33	90.5	3.6	2016	10	US-09-840-125-4	Sequence 4, Appli
34	90.5	3.6	581	9	US-09-298-523B-56	Sequence 56, Appl
35	90	3.5	595	10	US-09-853-033-2	Sequence 2, Appli
36	90	3.5	619	9	US-09-882-774-1	Sequence 1, Appli
37	90	3.5	807	9	US-10-044-692-5	Sequence 5, Appli
38	90	3.5	807	9	US-10-044-539-5	Sequence 5, Appli
39	90	3.5	2507	9	US-09-819-104A-2	Sequence 2, Appli
40	90	3.5	471	10	US-09-815-242-14028	Sequence 14028, A
41	89.5	3.5	660	10	US-09-853-033-6	Sequence 6, Appli
42	89.5	3.5	1399	9	US-09-388-221-4	Sequence 4, Appli
43	89.5	3.5	1424	9	US-09-388-221-12	Sequence 12, Appl
44	89.5	3.5	1429	9	US-10-028-392-11	Sequence 11, Appl
45	89.5	3.5				

ALIGNMENTS

RESULT 1

US-09-826-581-6

; Sequence 6, Application US/09826581

; Patent No. US20020142310A1

; GENERAL INFORMATION:

; APPLICANT: Andersson, Leif

; APPLICANT: Luthman, L. Holger

; APPLICANT: Marklund, Stefan

; TITLE OF INVENTION: VARIANTS OF THE HUMAN AMP-ACTIVATED PROTEIN KINASE GAMMA 3

; FILE REFERENCE: 11145-007001

; CURRENT APPLICATION NUMBER: US/09/826,581

; CURRENT FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: US 60/195,665

; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 489

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-826-581-6

Query Match	100.08;	Score	2538;	DB	10;	Length	489;
Best Local Similarity	100.08;	Pred. No.	1.5e-209;				
Matches	489;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MEPGLHALRRTPSWSSLGSGHQSFWSPAVTSSSRIRGKRAKALRW	60				
Db	1	MEPGLHALRRTPSWSSLGSGHQSFWSPAVTSSSRIRGKRAKALRW	60				
QY	61	TRKQSVEEGEPGQEGPRSPRAESTGLTATFPKTTPLAQADPAGVGTPTGDCPLPSD	120				
Db	61	TRKQSVEEGEPGQEGPRSPRAESTGLTATFPKTTPLAQADPAGVGTPTGDCPLPSD	120				
QY	121	CTASAGSTDDVELATFEFATEWECLEGLLEERPALCLSPQAPFPKLGWDELKRP	180				
Db	121	CTASAGSTDDVELATFEFATEWECLEGLLEERPALCLSPQAPFPKLGWDELKRP	180				
QY	181	AQIMRFMOEHTCYDAMATSSKLVIFDTMLEIKKFAFFALVANGYRAAPLWDSKKQSVGM	240				
Db	181	AQIMRFMOEHTCYDAMATSSKLVIFDTMLEIKKFAFFALVANGYRAAPLWDSKKQSVGM	240				

QY 241 LTITDFILVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLVSI SPNDSLFEAVYTL 300  
Db 241 LTITDFILVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLVSI SPNDSLFEAVYTL 300  
QY 301 IKNRHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPRPSFLYRTTODLIGITFRDLA 360  
Db 301 IKNRHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPRPSFLYRTTODLIGITFRDLA 360  
QY 361 VVLETAPILTALDIFVDRRVSAIPVNVNCGQVGLYSRFDVTHLAAQQTYNHLDMSVGEA 420  
Db 361 VVLETAPILTALDIFVDRRVSAIPVNVNCGQVGLYSRFDVTHLAAQQTYNHLDMSVGEA 420  
QY 421 LRQFTLCLEGLVSCQPHEISGEVIDRIAREQVHRLVLDVTHLGGVWSLSLIDLOALVLS 480  
Db 421 LRQFTLCLEGLVSCQPHEISGEVIDRIAREQVHRLVLDVTHLGGVWSLSLIDLOALVLS 480  
QY 481 PAGIDALGA 489  
Db 481 PAGIDALGA 489

## RESULT 2

US-09-925-297-461  
; Sequence 461, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 461  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-461

Query Match 40.1%; Score 1017; DB 10; Length 344;  
Best Local Similarity 64.8%; Pred. No. 3.6e-79;  
Matches 193; Conservative 53; Mismatches 52; Indels 0; Gaps 0;  
QY 183 IYMRPMQEHYCYDAMATSSKLVIFUTMLEIKKAFALVANGVRAAPLWDSKKQSFVGM 242  
Db 41 VYTFMKSHRCYDLIPTSSKLVVFTSLQVKKAFALVTNGVRAAPLWDSKKQSFVGM 100  
QY 243 ITDFTLVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLVSI SPNDSLFEAVYTLIK 302  
Db 101 ITDFTLVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLVSI SPNDSLFEAVYTLIK 160  
QY 303 NRIHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPRPSFLYRTTODLIGITFRDLAV 362  
Db 161 NKIHLRPVDPESGNTLYLTHKRLKFLHIFGSLPRPSFLYRTTODLIGITFRDLAV 220  
QY 363 LETAPILTALDIFVDRRVSAIPVNVNCGQVGLYSRFDVTHLAAQQTYNHLDMSVGEALR 422  
Db 221 RYTFVYVGLGIFVQHRVSAIPVNVNCGQVGLYSRFDVTHLAAQQTYNHLDMSVGEALR 280  
QY 423 QRTLCLEGLVSCQPHEISGEVIDRIAREQVHRLVLDVTHLGGVWSLSLIDLOALVLS 480  
Db 281 HRSHYFEGVLKCYLHETLETINRLVVEAEVHRLVVVDENVKGVISLSLIDLOALVLT 338

## RESULT 3

US-10-108-605-71  
; Sequence 71, Application US/10108605  
; Patent No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCO  
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCO  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108,605  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 71  
; LENGTH: 1207  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-71

Query Match 35.2%; Score 892.5; DB 9; Length 1207;  
Best Local Similarity 40.6%; Pred. No. 1.1e-67;  
Matches 202; Conservative 93; Mismatches 121; Indels 81; Gaps 10;  
QY 7 HALRTPSWSSLGSGHEQMSFLEQENSSSWPSPAVTS-----SSER 48  
Db 594 HSLGHSPSAAAAAQQTLVYERRRVSSAHSSPSTRHSPVHQCLMRRRSYSVEQ 653  
QY 49 IRGKRKAKALRWTRKSVBEGPPGQGGPRSRPAESTGLEATFPKTPPLAQADPAGVG 108  
Db 654 IE-----QWKQQQLL-----GRKISLHFYSNHAGRLVLGGG 688  
QY 109 T-PPTGWDCLPSDCTASAGSSTDDVELATEF-PATEAWCELEGLEBERPALCLSPAP 166  
Db 689 NRRPTDVLCP-----LESIKEFEQISANKRLLAELKKEE----- 725  
QY 167 FPKLGMDELKPGAQIYMRPMQEHYCYDAMATSSKLVIFUTMLEIKKAFALVANGVRA 226  
Db 726 -----DD-----SQIFVKFRFKCYDLIPTSAKLWFDLTLLVKAFYALVYNGVRA 773  
QY 227 APWDSKKQSFVGMILTITDFTLVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLV 286  
Db 774 APWDSKKQSFVGMILTITDFTLVHRYRSPVQVIEIHOHKIETWREIYLOGCFKPLV 832  
QY 287 ISPNDLSFEAVYTLIKNRHRLPVLDPVSGNVLIHLTHKRLKFLHIFGSLPRPSFLYR 346  
Db 833 ICPDASLDAIKLILHSRHLRPVDPATGNVLIHLTHKRLKFLHIFGSLPRPSFLYR 892  
QY 347 TIQDLIGITFRDLAVLETAIPILTALDIFVDRRVSAIPVNVNCGQVGLYSRFDVTHLAA 406  
Db 893 SLRELKIGTYNNIETADETTSITALKKFVERRVSAIPVNVNCGQVGLYSRFDVTHLAA 952  
QY 407 QQTYNHLDMSVGEALRQRTLCLEGLVSCQPHEISGEVIDRIAREQVHRLVLDVTHLGG 466  
Db 953 ERTYNLDVSLRKANEHRNFEQVKCNLDSELTIMERIVRAEYHRLVVVDENVKGVIS 1012  
QY 467 VVSLSDILQALVLS 483  
Db 1013 IISLSDILQLVLRPSG 1029

## RESULT 4

US-10-108-605-195  
; Sequence 195, Application US/10108605  
; Patent No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCO  
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

```
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 195
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-195

Query Match          4.5%; Score 115; DB 9; Length 980;
Best Local Similarity 22.5%; Pred. No. 0.49;
Matches 103; Conservative 66; Mismatches 187; Indels 102; Gaps 22;

QY 85 ESTGLEATFPKTTPLAQADPAGVGTPTGWDCLPSDCTASAGSSTDD--VELATEFFAT 142
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 87 EKAGDEIWEIKICQVCKDPS-----VEITQDIVESSEDERFEDMSDTAPPI 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 143 EAWEELEGL--LEERPAICLSQAPPPKLGWDDDLRKPQAQIYMRMOE--HTCYDAMA 198
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 134 ELPPCELSRLEDISETIQSCSLSTPLRKEKLSMALE-----SEYIKKLLNLFHVCELDN 188
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 199 TSSKLVIFDTMLEIKKAPFALVANGYRAAPLWDSKKQSFVCMLTITDFIL--VLHRYYSR 256
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 189 TEGLHHLF-----EIFKNIFLNKNALFEIMFADDTIFDVGCLEYDPSVQPKKRQYLK 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 257 PLVQIVEIEQHK-----IETWREIYLCQCFKPLVSI PNDSLFEAVYTLI--KNRIHR 307
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 245 QLAKEFRAVPKINLDDLAKIHQTFVQYIQDIIILTPSVFVEDNMLNLTSSFFFNKVEI 304
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 308 LPVL-----DPVSGNVHLHTHKRLKPLHIFGSLLP--RPSFLYRTIQD 350
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 305 VTMIQDERYLLDVFVAVLTDPITGDAKRRTDVLFLKPCNYAQNLOPQCKDSF-YKLTLC 363
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 351 LGICFTRDLAVLE-----TAPILTALDIEVDRRVSALPVYNCEQOVGLYSRFDV-- 401
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 364 LGLIQAELTILVMDKTKSASIDILTAIVEF-----SPLVVRN---YTLNQRANRPEVER 415
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 402 --IHLLAAQOTNHLDMVSGFALRORTLCLEGVLS--CQPHESL---GEVIDRIARQVHR 454
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 416 MLLNTAIEQLMNDSEPELGIAVQ-----LMGIVKILLEPENMLTEKGFNLNFYKYSVQT 470
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 455 LV-----LVD-----FTQHLIGVWSLSLTLQALV 478
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 471 LVAPVILNTIGDRPNQEDYQTAGLLGIV--LDILSFCV 506
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-864-761-33979
; Sequence 33979, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
```

```
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33979
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006966.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.9
; OTHER INFORMATION: SWISSPROT HIT: P80385, EVALUATE 2.00e-05
; OTHER INFORMATION: EST_HUMAN HIT: BE763984.1, EVALUATE 2.00e-15
; OTHER INFORMATION: EST_HUMAN HIT: AW957020.1, EVALUATE 1.00e-09
US-09-864-761-33979

Query Match          4.5%; Score 114; DB 10; Length 91;
Best Local Similarity 65.6%; Pred. No. 0.018;
Matches 21; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 452 VHRLLVDETFQHLGWWSLSLTLQALVSPAG 483
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 VHRLLVVVNEADSIIVGIISLILQALILTPAG 32
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-10-001-873-50
; Sequence 50, Application US/10001873
; Patent No. US20020160388A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and
```

```
; FILE REFERENCE: DEX-0275
; CURRENT APPLICATION NUMBER: US/10/001.873
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,055
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252,496
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1134
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-873-50

Query Match      4.2%; Score 105.5; DB 9; Length 1134;
Best Local Similarity 23.3%; Pred. No. 4;
Matches 70; Conservative 46; Mismatches 113; Indels 71; Gaps 13;

Qy 38 PPAVTSSEIRIRKRAKALRWTRQKSVVEGPPGGGPRSR-----PAAEST 87
      |||| | | | | | | | | | | | | | | | | | | | | | |
Db 68 PSPAGCGG---GLLEAQAQSATQSCAEPSECPDFVEGPEPRVDSRTEPCATA 123
      |||| | | | | | | | | | | | | | | | | | | | | | |

Qy 88 GLEATFKTTPLAQADP--AGVCTPTGTGWDCLPSDCTASAGSSTDVELATEFP 145
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 124 GVQLT-PETLVEAKEPEFVPGVPV--EAVPEGLAQVAPSEQ-----PT 172
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 146 ECEGLELIERPALCLSPAPPKLGWDEDLKPKGAQIYMRPMOHTC-YDAMATSS 203
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 173 DCDVPAGEGQCP--LEPQEAIVLGSTCFLEERASSDQFLPSLEDPAGMNA 230
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 204 -----VIFDTMLEIKKAFALVANGVRAAPLWDSKKQSPVGMGITITFIL 254
      || | | | | | | | | | | | | | | | | | | | | | |
Db 231 QARPLSPGAAGAQAQALEAAESLVL-----EQSFLHGITL----- 267
      || | | | | | | | | | | | | | | | | | | | | | |

Qy 255 RSPLVQIYEIQHKIETWREIYLGCFKPLYSISPNDLSFEAVYTLIKNRIH 314
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 268 ---LSEIAELELER----RSQMGGAERLVARPSLESLLAGSHMLREVLDG 319
      |::| | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-975-719-287
; Sequence 287, Application US/09975719
; Publication No. US20030022349A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Rahme, Laurence G.
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361003
; CURRENT APPLICATION NUMBER: US/09/975,719
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-287

Query Match      4.0%; Score 101; DB 9; Length 187;
Best Local Similarity 25.0%; Pred. No. 0.68;
Matches 44; Conservative 14; Mismatches 46; Indels 72; Gaps 10;

Qy 38 PPAVTSS-----SERINGKRAKALEWTRQ-----KSVEEGPPGQG 75
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 19 PAPCTTRSCPSPRSGRGTSRAGSDR-RGRRANGARRWTRRPPRGRSLADAAP 77
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 76 EGFRSRPAEST-CLEAT-----FPKTTPLAQAD-----PAGVCTPPTG 116
      |::| | | | | | | | | | | | | | | | | | | | | |
```

```
Db 78 AASRARPRASSTFGYRSTWKPLRRPRSPCCAAATGSRSCRAGRPSAGCVPGP 137
      |::| | | | | | | | | | | | | | | | | | | | | |
Qy 117 LPFD-----CTASAAGSSTDVELATEFPATEWBECEGLEELERPALCLSP 168
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 138 APCSRGRPCAAP-----PPSPAW-----PARSAGSASP 167
      |::| | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-738-626-6006
; Sequence 6006, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6006
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6006

Query Match      4.0%; Score 100.5; DB 9; Length 440;
Best Local Similarity 19.8%; Pred. No. 2.6; Mismatches 62; Indels 67; Gaps 15;
Matches 64; Conservative 62; Mismatches 130; Indels 67; Gaps 15;

Qy 180 GQAIYRMFMOEHTCYDAMATSSKLVIFDTMLEIKKAFALVANGVRAAPL 239
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 47 GSASLLRVIDERALHNM-----LIMLRTLDDASAAVFA---GALAVN 97
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 240 MLTITDFT--LVLHRYRSPVQIYEIQHKIETWREIYLGCFK---PLYS 286
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 98 AIVVVSLLTFAVGVFGRT-----VGRKNPYSVMLRSVAVLSGLAKTL 153
      |::| | | | | | | | | | | | | | | | | | | | | |
Qy 287 ISPNDLSFEAVYTLIKNRIHRLPVLDPVSGNVLHILTKLLAFHIFGS----- 340
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 154 IAPGPGFNGPYA---TEVELRMVDIAQEHGIVEIERMIQSVFDLASTIV 210
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 341 PSFLYRTIQDLIGITFRDLAWVLETAPILTALDIFVDRRVSPALPVYNE-- 399
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 211 PEMIW-----IESG-----KTAGQATAL--CVRSRHSRIPVIGENVD 254
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 400 DVIHLAAQOTYNHLD-----MSVGEALRQTLCLSGVLSQCPHESLGEY 455
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 255 DLV-----QKTYATDGGKSVLDEVMRBATFVPOS-----KSDALL 303
      |::| | | | | | | | | | | | | | | | | | | | | |

Qy 456 VLVDETOHLLGVWSLSDILQALV 478
      |::| | | | | | | | | | | | | | | | | | | | | |
Db 304 ILVDEYGVAGLISIEDILEEIV 326
      |::| | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-09-272-809-2
; Sequence 2, Application US/09272809
```

```

; Patent No. US200200222329A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Lagarias, John
; TITLE OF INVENTION: phyto
;
; FILE REFERENCE: 2500.118U
;
; CURRENT APPLICATION NUMBER:
;
; CURRENT FILING DATE: 1998
;
; NUMBER OF SEQ ID NOS: 24
;
; SOFTWARE: Patentin Ver. 2
;
; SEQ ID NO 2
;
; LENGTH: 1276
;
; TYPE: prt
;
; ORGANISM: Unknown
;
; FEATURE:
;
; OTHER INFORMATION: Descr
US-07-272-809-2

```

Query Match	3.9%;	Score 100;	DB 10;	Length 1276;
Best Local Similarity	20.3%;	Pred. No. 14;		
Matches 115;	Conservative 72;	Mismatches 165;	Indels 214;	Gaps 34;
QY	79	RSRPAA--ESTGLEATFPKTTTLLAQADPAGVGTPTGWCDCLP-----SDCTAS	124	
Db	276	RGQFSRAMEPMAAVQSTWEXPRFTSVAP-----LPPT--NCVPHGYTLGELQSGDWIAP	329	
QY	125	AAGSSTDQVE--LATERPATEAMECELEGLLEER-----PALCLSPQ	164	
Db	330	PESLSAENFQSELIIVPLAAQQWVGSLLILLREKESLVKHWAGKRGIDRRNILPRLSFEAW	389	
QY	165	APPFKL--CWDEELRK-----PQAQIYM-----RFMOEHTCVDAMATSSKLIVFTDML	210	
Db	390	EETQKILPWNNSERKLAQVASTQIYNATQOFTVRLITQCTAYDPLTQLPNNIIFNRQL	449	
QY	211	EIKKAZFALVANGVRAAPLMDSKQSFVGMLTITDILVLRHYR-----	255	
Db	450	TL-----ALL-----DALYEGK---MGVGL-----VIAMDRPKRINESGHKTGDGLQ	490	
QY	256	-----SPLY-----QIYETEQUHKIETWRE-IYLOQ	279	
Db	491	EVADRNLNQKLSPLAAYSPLLSRWHGDGFTILLTQISDNQEMIPLCERLLSTQEPFFLQ	550	
QY	280	CFKPL-----VSTSPNDLSFEAVYTLK-----NRHR-----LPVLD	312	
Db	551	--QPIVLTASMGISTAPYDG--ETAESLLKFAEIALTRAKCQGNKNTQYFRQDSAPMLD	606	
QY	313	--PVSGNVLIITHKRLIKLFIHFCSLPLPMISFLYKTIQDLGIGCTGFEDLAWVLETAPIL--	369	
Db	607	RLTLESDLRQALTNQEFVLVE-----QPQVALITCKLIGVE-----ALVRWQHPIRLG	653	
QY	370	-TALDIFVDRRYSALPVVNYECQGVVGLYSKFQVILHAAQOTYNIHLDMSVCEALR-----	422	
Db	654	QVAPDVFIP-LAELEGILNHIGOW-----LETACATHQHFRETGRRLRMVAVNIS	703	
QY	423	---QRTLCLEGLVSC---QPHEISLG-EVIDRLAREQV-----HRL-----VLVD	459	
Db	704	ARQFQDEKWLNSVLBCLARTCMPDEDLELITESLMMEDIKGTVVLHLRREEGVQVAID	763	
QY	460	ETOHLLGVVSLSDILOALVSPAGID	485	
Db	764	DFG--TCYSSLS-ILKOLPIHRLIKD	786	

```

; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 701
; TYPE: PRT
; ORGANISM: human
US-10-052-092-12

Query Match          3.9%; Score 99.5; DB 9; Length 701;
Best Local Similarity 17.2%; Pred. No. 6.4;
Matches 81; Conservative 66; Mismatches 126; Indels 197; Gaps 22;

QY 37 WPSAVTSSSEIRKRAKALRWTKQSVSVE--RGEPPGGEGPSRPAARSTGLEATF 93
      |||| : : : || : : : || : : : || : : : || : : : || : : :
Db 292 WPSPLMK-----RSKKNSLALSITADQWVSALLDAEPP-----ILYSEY 331
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 94 PKTTPLAQADPAGVCTPPTGNDCLPSDCTASAAGSSDDVELATEFPATEAMECELEGL 153
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 332 DPTRPSEASMGLLT-----NLADRELVIHI-----NNAKRVPGEV 368
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 154 EERPALCSQAPFPKLGWDELKPKQAQIYMRFOEHTCYDAMATSSKLIVDTMLEIK 213
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 D----LTLHDQVHLLCAW-----LEILMIGLVWRSMEHPVKLLFAPNLL--- 409
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 214 KAFALVANGVRAAPLWDSKQSVGMLTITPILVHLHYRSPVLVOIYIEBOHKIETWR 273
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 -----LDRNQGKVEGVMEIFDMLLATSSRFM----- 437
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 274 EYILOG---CFKLVSVISPNDSLFEAVYTLI-----KNRIHRLPYLDPVSGNVLI 320
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 -WNLOGEEFVCLKSILIN-----SCVYTFISSTLKSLEKDIHR--VLDKITDTLIH 488
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 ILT-----HKRLKFLHIFGSLLPSPFLYRTIIDLIGITFRDLAYVLETAPILTA 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 LMAKAGLTQQQHQRLAQLLLSHI-----RHMNRQG----- 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 372 LDIFVDRVSALPVYNECGVGVGLYRSPDVHILAAQOTYINHLDMXSGEALRQRTLCLEGV 431
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 522 -----KCV-VEGVMEIFDKL-LATSSRFMMNLQ-GEEF---VCLKSI 557
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 432 L-----SCQPHESLGEVIDRIAREOVHRLV---LVDETQH 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 558 ILLNSGVYTFSLSTLKSLEEKDTHRVLDKAITDTLIHLMAKAGLTQQQH 607
      : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-211-613-1
; Sequence 1, Application US/10211613
; Publication No. US20030099641A1
; GENERAL INFORMATION:
; APPLICANT: CELL SIGNALING TECHNOLOGY, INC.
; APPLICANT: SMITH, Bradley
; APPLICANT: CROSBY, Katherine
; APPLICANT: WU, Jiong
; TITLE OF INVENTION: Monoclonal Antibodies Specific for Phosphorylated Estrone
; FILE REFERENCE: CST-202
; CURRENT APPLICATION NUMBER: US/10/211,613
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/310,066
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Homo sapiens

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:55:08 ; Search time 6.60595 Seconds  
(without alignments)  
827.023 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91  
Perfect score: 214  
Sequence: 1 GKRRAKALRWTRQKSVSEGE.....PQGGGPRSRPAESTGLEA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	214	100.0	489 22 AAB47679 PRKAG3. Homo sapi
2	210	98.1	464 22 AAE00223 Human AMPK gamma s
3	144	67.3	464 23 AAE22985 Pig PRKAG3 polymor
4	142	66.4	464 22 AAE00222 Pig AMPK gamma sub
5	142	66.4	464 23 AAE22984 Pig wild-type PRKA
6	142	66.4	464 23 AAE22987 Pig PRKAG3 polymor
7	142	66.4	464 23 AAE22988 Pig PRKAG3 polymor
8	142	66.4	514 22 AAE00224 Sus scrofa Prkag3
9	136	63.6	464 23 AAE22986 Pig PRKAG3 polymor
10	64.5	30.1	171 22 ABG03666 Novel human diagno

11	63.5	29.7	61	21	AAG35967	Zea mays protein f
12	61.5	28.7	415	19	AAW71562	Human hepatocyte n
13	61.5	28.7	630	19	AAW71560	Human hepatocyte n
14	61.5	28.7	631	19	AAW71580	Human hepatocyte n
15	61.5	28.7	631	19	AAW71559	Human native hepat
16	61.5	28.7	866	22	ABB67322	Drosophila melanog
17	61.5	28.7	933	22	ABB67321	Drosophila melanog
18	61.5	28.7	958	22	ABB58729	Drosophila melanog
19	60.5	28.3	152	22	ABG26522	Novel human diagno
20	60	28.0	73	22	AAU41694	Propionibacterium
21	60	28.0	187	20	AAV29189	Amino acid sequenc
22	60	28.0	270	20	AAV29328	Human secreted pro
23	60	28.0	270	22	AAU39054	Human secreted pro
24	60	28.0	270	23	ABB55763	Human polypeptide
25	59.5	27.8	1221	20	AAV05940	Thermophilus therm
26	59.5	27.8	1255	20	AAV05944	Thermophilus therm
27	59	27.6	230	22	ABG08090	Novel human diagno
28	59	27.6	705	22	AAV39328	Human polypeptide
29	59	27.6	714	22	AAV41114	Human polypeptide
30	59	27.6	863	22	ABB64979	Drosophila melanog
31	58	27.1	144	22	ABG03531	Novel human diagno
32	58	27.1	271	23	AAU96713	Human osteoclast a
33	58	27.1	282	23	AAU96712	Human osteoclast a
34	56.5	26.4	529	21	AAV83085	F-box protein FBP-
35	56.5	26.4	529	22	AAB48307	Human p26 protein
36	56	26.2	405	23	AAE14452	Human protein phos
37	56	26.2	447	23	AAU82751	Amino acid sequenc
38	56	26.2	449	23	ABG27792	Novel human diagno
39	56	26.2	493	23	AAU75786	Human protein phos
40	56	26.2	1133	22	AAE04836	Human SGP018 phosp
41	55.5	25.9	102	22	AAO10686	Human polypeptide
42	55.5	25.9	107	22	AAG92689	C glutamic prote
43	55.5	25.9	345	21	AAV70506	Human BAG-1L prote
44	55.5	25.9	601	22	ABG23745	Novel human diagno
45	55.5	25.9	678	22	AAV25801	Human protein sequ

ALIGNMENTS

RESULT 1			
AAB47679			
ID	AAB47679	standard; Protein; 489 AA.	
AC	AAB47679;		
XX			
XX			
DT	21-JAN-2002	(first entry)	
XX			
DE	PRKAG3.		
XX			
KW	Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;		
KW	metabolic disease; diabetes; obesity; substitution; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 71	/note= "Possible variation point P71A"	
FT	Misc-difference 340	/note= "Possible variation point R340W"	
XX			
PN	WO200177305-A2.		
XX			
PD	18-OCT-2001.		
XX			
PF	06-APR-2001; 2001WO-SE00765.		
XX			
PR	07-APR-2000; 2000US-195665P.		
XX			
PA	(AREX-) AREXIS AB.		
XX			
PI	Andersson L, Luthman H, Marklund S;		
XX			



DR WPI; 2001-657170/75.  
 DR N-PSDB; AAH43685.  
 XX  
 PT New variants of human AMP-activated protein kinase gamma3 subunit  
 PT associated with a metabolic disease e.g. diabetes or obesity and method  
 PT for determining a risk estimate of diseases in subject by detecting the  
 PT variant -  
 XX  
 PS Disclosure; Fig 5; 25pp; English.  
 XX  
 CC This sequence is encoded by the full length cDNA encoding the human  
 CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting  
 CC the presence of the PRKAG3 DNA, or a variant, is useful in determining  
 CC a risk estimate of a metabolic disease, such as diabetes or obesity,  
 CC in a subject. The variation may occur in exons 3, 4 or 10. In exon  
 CC 3 variation may be a substitution of a G for a C at nucleotide 320,  
 CC resulting in the amino acid substitution P71A; in exon 4 variation may  
 CC be a substitution of a T for a C at nucleotide 550; and in exon 10  
 CC variation may be a substitution of a T for a C at nucleotide 1037,  
 CC resulting in the amino acid substitution R340W. There may also be  
 CC nucleotide variation in intron 6. The numbering of these  
 CC variations is based on the full length cDNA, rather than on  
 CC position 1 of the open reading frame.  
 XX  
 SQ Sequence 489 AA;  
 Query Match 100.0%; Score 214; DB 22; Length 489;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-19;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVGEPPGPGGEGPRSPRAESTGLEA 41  
 DB 51 GKRRKALRWTRQKSVGEPPGPGGEGPRSPRAESTGLEA 91  
 RESULT 2  
 AAE00223  
 ID AAE000223 standard; Protein; 464 AA.  
 XX  
 AC AAE00223;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
 XX  
 KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiac; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 172..225  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Misc-difference 200  
 FT /note= "RN- mutation site"  
 FT Domain 253..307  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 329..382  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 400..453  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 XX  
 PN WO200120003-A2.  
 XX  
 PD 22-MAR-2001.  
 XX  
 PF 11-SEP-2000; 2000WO-EP09896.

XX  
 PR 10-SEP-1999; 99EP-0402236.  
 PR 18-MAY-2000; 2000EP-0401388.  
 XX  
 PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Geillin J, Le Roy P, Chardon P;  
 XX  
 DR WPI; 2001-244810/25.  
 DR N-PSDB; AAD03320.  
 XX  
 PT New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 PS Claim 5; Fig 3; 71pp; English.  
 XX  
 CC The present sequence is human adenosine monophosphate  
 CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
 CC complete PRKAG3. Mutation in prkag3 results in an altered regulation of  
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
 CC useful as therapeutic for treating carbohydrate metabolism disorders such  
 CC as diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 XX  
 SQ Sequence 464 AA;  
 Query Match 98.1%; Score 210; DB 22; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 1.9e-18;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVGEPPGPGGEGPRSPRAESTGLEA 41  
 DB 26 GKRRKALRWTRQKSVGEPPGPGGEGPRSPRAESTGLEA 66  
 RESULT 3  
 AAE22985  
 ID AAE22985 standard; Protein; 464 AA.  
 XX  
 AC AAE22985;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig PRKAG3 polymorphic variant (PRKAG3-30).  
 XX  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 30 /note= "Wild type Asn is substituted with Thr due

FT XX to single nucleotide polymorphism (SNP) "

PN W0200220850-A2.

XX 14-MAR-2002.

XX 10-SEP-2001; 2001WO-0528283.

XX 08-SEP-2000; 2000US-231045P.

PR 08-JAN-2001; 2001US-260239P.

PR 18-JUN-2001; 2001US-299111P.

XX (IOWA ) UNIV IOWA STATE RES FOUND INC.

PA Rothschild MF, Ciobanu DC, Malek M, Plastow G;

XX WPI: 2002-392850/42.

XX N-PSDB; AAD36457.

XX Screening animals to determine those likely to produce larger litters

PT and improved meat quality traits involves assaying for the presence of

PT polymorphisms in the AMP activated protein kinase regulatory gamma

PT subunit gene -

XX Disclosure; Page 91-93; 109pp; English.

XX The invention relates to a method for screening animals to determine

CC those more likely to produce large litters and improved meat quality

CC traits. The method involves assaying for the presence of a genotype

CC in the sample of genetic material obtained from animal. The genotype

CC is characterised by polymorphism(s) in the AMP activated protein

CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used

CC for screening animals e.g., pigs to determine those most likely to

CC exhibit improved meat quality traits and to produce larger litters.

CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-30).

XX

SQ Sequence 464 AA;

Query Match 67.3%; Score 144; DB 23; Length 464;

Best Local Similarity 68.3%; Pred. No. 4.7e-10;

Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GKRRAKALRWTRKSVGEPPGCGGPRSRPAESTGLEA 41

1 : || |||||: |||| ||| ||||| ||||| ||

26 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 66

Db

RESULT 4

AAE00222

ID AAE00222 standard; Protein: 464 AA.

XX

AC AAE00222;

XX

DT 13-JUN-2001 (first entry)

XX

DE Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.

XX

KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;

KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;

KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;

KW cystathione beta synthase; CBS; cardiatic; gene therapy; RN locus;

KW chromosome 15.

OS Sus scrofa.

XX

XX

FH Key Location/Qualifiers

FT Domain 172..225

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Misc-difference 200

FT /note= "RN- mutation site"

FT Domain 253..307

FT /label= CBS

FT

FT Domain /note= "Cystathione beta synthase domain"

FT 329..382

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT 400..453

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

XX

XX W0200120003-A2.

PN 22-MAR-2001.

XX

XX 11-SEP-2000; 2000WO-EP09896.

XX

XX 10-SEP-1999; 99EP-0402236.

PR 18-MAY-2000; 2000EP-0401388.

XX

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA (ANDE/) ANDERSSON L.

PA (LOOF/) LOOFT C.

PA (KALM/) KALM E.

XX

XX Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI Tannuicelli N, Gellin J, Le Roy P, Chardon P;

PI N-PSDB; AAD03319.

DR WPI: 2001-244810/25.

XX

XX New variants of the gamma subunit of vertebrate adenosine

PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and

PT myopathy -

XX

PS Claim 5; Fig 3; 71pp; English.

XX

CC The present sequence is pig adenosine monophosphate

CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,

CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome

CC 15. Mutation in prkag3 results in an altered regulation of carbohydrate

CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as

CC therapeutic for treating carbohydrate metabolism disorders such as

CC diabetes, obesity, and disorders associated with muscle metabolism

CC such as myopathy and cardiovascular diseases, to modulate AMPK

CC activity, and for restoring a normal AMPK function. PRKAG3 sequence

CC and its functionally altered mutants are useful for the diagnostic

CC evaluation, genetic testing and prognosis of a metabolic disorder,

CC preferably a carbohydrate metabolism disorder. Primers that can detect

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

CC useful for detecting a dysfunction of carbohydrate metabolism resulting

CC from the expression of a functionally altered allele of PRKAG3.

CC Transgenic animal and host cell transformed with PRKAG3 or a

CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

CC screening compounds able to modulate AMPK activity. Nucleic acid

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or

CC in a sequence encoding the first cystathione beta synthase (CBS) domain

CC of PRKAG3 and is useful in gene therapy.

XX

SQ Sequence 464 AA;

Query Match 66.4%; Score 142; DB 22; Length 464;

Best Local Similarity 68.3%; Pred. No. 8.4e-10;

Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GKRRAKALRWTRKSVGEPPGCGGPRSRPAESTGLEA 41

1 : || |||||: |||| ||| ||||| ||||| ||

26 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 66

Db

RESULT 5

AAE22984

ID AAE22984 standard; Protein: 464 AA.

XX

XX AAE22984;

AC

```
XX 09-AUG-2002 (first entry)
XX Pig wild-type PRKAG3 protein.
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
XX screening; meat quality; single nucleotide polymorphism; SNP; pig.
XX Sus scrofa.
XX Key Location/Qualifiers
FH Misc-difference 30 /note= "Wild type Asn is replaced with Thr during
FT single nucleotide polymorphism (SNP)"
FT Misc-difference 52 /note= "Wild type Gly is replaced with Ser during
FT single nucleotide polymorphism (SNP)"
FT Misc-difference 199 /note= "Wild type Val is replaced with Ile during
FT single nucleotide polymorphism (SNP)"
FT Misc-difference 200 /note= "Wild type Arg is replaced with Gln during
FT single nucleotide polymorphism (SNP)"
XX WO200220850-A2.
XX 14-MAR-2002.
XX 10-SEP-2001; 2001WO-US28283.
XX 08-SEP-2000; 2000US-231045P.
XX 08-JAN-2001; 2001US-260239P.
XX 18-JUN-2001; 2001US-299111P.
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX WPI; 2002-393850/42.
XX N-PSDB; AAD36456.
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene .
XX Claim 2; Fig 1; 109pp; English.
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig wild-type PRKAG3 protein.
XX Sequence 464 AA;
XX Query Match 66.4%; Score 142; DB 23; Length 464;
XX Best Local Similarity 68.3%; Pred. No. 8.4e-10;
XX Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
XX
XX 1 GKRRKALRWTRQKSVGEPPGGGEGPRSRPAAESTGLEA 41
XX I : || |||||: ||||| ||| |||||: ||||| ||
XX 26 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 66
XX
XX RESULT 6
XX AAE22987
XX ID AAE22987 standard; Protein; 464 AA.
XX
```

```
AC AAE22987;
XX 09-AUG-2002 (first entry)
XX Pig PRKAG3 polymorphic variant (PRKAG3-199).
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
XX screening; meat quality; single nucleotide polymorphism; SNP; pig;
XX variant.
XX Sus scrofa.
XX Key Location/Qualifiers
FH Misc-difference 199 /note= "Wild type Val is substituted with Ile due
FT to single nucleotide polymorphism (SNP)"
FT Misc-difference 200 /note= "Wild type Val is substituted with Ile due
FT to single nucleotide polymorphism (SNP)"
XX WO200220850-A2.
XX 14-MAR-2002.
XX 10-SEP-2001; 2001WO-US28283.
XX 08-SEP-2000; 2000US-231045P.
XX 08-JAN-2001; 2001US-260239P.
XX 18-JUN-2001; 2001US-299111P.
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX WPI; 2002-393850/42.
XX N-PSDB; AAD36459.
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene .
XX Disclosure; Page 100-102; 109pp; English.
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterised by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-199).
XX Sequence 464 AA;
XX Query Match 66.4%; Score 142; DB 23; Length 464;
XX Best Local Similarity 68.3%; Pred. No. 8.4e-10;
XX Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
XX
XX 1 GKRRKALRWTRQKSVGEPPGGGEGPRSRPAAESTGLEA 41
XX I : || |||||: ||||| ||| |||||: ||||| ||
XX 26 GDQGNKASRWTRQEDVEGGPPGREGPQSRPVAESTGQEA 66
XX
XX RESULT 7
XX AAE22988
XX ID AAE22988 standard; Protein; 464 AA.
XX
XX AAE22988;
XX 09-AUG-2002 (first entry)
XX Pig PRKAG3 polymorphic variant (PRKAG3-200).
XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;
```

KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 OS Sus scrofa.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 200  
 FT /note= "Wild type Arg is substituted with Gln due  
 FT to single nucleotide polymorphism (SNP)"  
 FT  
 XX  
 XX W0200220850-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 XX 10-SEP-2001; 2001WO-US28283.  
 XX  
 XX 08-SEP-2000; 2000US-231045P.  
 PR 08-JAN-2001; 2001US-260239P.  
 PR 18-JUN-2001; 2001US-299111P.  
 XX  
 XX (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 PA  
 XX Rothschild WF, Ciobanu DC, Malek M, Piastow G;  
 XX  
 XX WPI: 2002-393850/42.  
 DR N-PSDB; AAD36460.  
 DR  
 XX Screening animals to determine those likely to produce larger litters  
 PT and improved meat quality traits involves assaying for the presence of  
 PT polymorphisms in the AMP activated protein kinase regulatory gamma  
 PT subunit gene -  
 XX  
 XX Disclosure; Page 105-107; 109pp; English.  
 XX  
 CC The invention relates to a method for screening animals to determine  
 CC those more likely to produce large litters and improved meat quality  
 CC traits. The method involves assaying for the presence of a genotype  
 CC in the sample of genetic material obtained from animal. The genotype  
 CC is characterised by polymorphism(s) in the AMP activated protein  
 CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
 CC for screening animals e.g., pigs to determine those most likely to  
 CC exhibit improved meat quality traits and to produce larger litters.  
 CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).  
 XX  
 SQ Sequence 464 AA;  
 Query Match 66.4%; Score 142; DB 23; Length 464;  
 Best Local Similarity 68.3%; Pred. No. 8.4e-10;  
 Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVVEEGPPGGGKPRSRPAESTGLEA 41  
 Db 26 GDGCKASRWTRQEDVEGGPPGPRGPGQSRPVAESTGQEA 66  
 RESULT 8  
 AAE00224  
 ID AAE00224 standard; Protein; 514 AA.  
 AC AAE00224;  
 XX  
 XX 13-JUN-2001 (first entry)  
 XX  
 XX Sus scrofa Prkag3 splice variant.  
 XX  
 XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiast; gene therapy.  
 XX  
 OS Sus scrofa.  
 XX  
 XX W0200120003-A2.  
 PN

XX 22-MAR-2001.  
 PD  
 XX  
 PF 11-SEP-2000; 2000WO-EP09896.  
 XX  
 XX 10-SEP-1999; 99EP-0402236.  
 PR 18-MAY-2000; 2000EP-0401388.  
 XX  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 XX Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX  
 XX WPI: 2001-244810/25.  
 DR N-PSDB; AAD03321.  
 XX  
 XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 XX Claim 5; Page 70-71; 71pp; English.  
 PS  
 XX  
 CC The present sequence is pig adenosine monophosphate (AMP)-activated  
 CC kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice  
 CC variant. Prkag3 gene is located in the RN locus of chromosome 15.  
 CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 XX  
 XX Sequence 514 AA;  
 SQ  
 Query Match 66.4%; Score 142; DB 22; Length 514;  
 Best Local Similarity 68.3%; Pred. No. 9.3e-10;  
 Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVVEEGPPGGGKPRSRPAESTGLEA 41  
 Db 76 GDGCKASRWTRQEDVEGGPPGPRGPGQSRPVAESTGQEA 116  
 RESULT 9  
 AAE22986  
 ID AAE22986 standard; Protein; 464 AA.  
 XX  
 XX AAE22986;  
 AC  
 XX  
 XX 09-AUG-2002 (first entry)  
 DT  
 XX Pig PRKAG3 polymorphic variant (PRKAG3-52).  
 DE  
 XX AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 XX

```

OS Sus scrofa.
XX Key Location/Qualifiers
FH Misc-difference 52
FT /note= "Wild Type Gly is substituted with Ser due
FT to single nucleotide polymorphism (SNP)"
XX
XX WO200220850-A2.
XX
XX 14-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US28283.
XX
XX 08-SEP-2000; 2000US-231045P.
XX
XX 08-JAN-2001; 2001US-260239P.
XX
XX 18-JUN-2001; 2001US-299111P.
XX
XX (IOWA ) UNIV IOWA STATE RES FOUND INC.
XX
XX Rothschild MF, Ciobanu DC, Malek M, Plastow G;
XX
XX WPI; 2002-393850/42.
XX
XX N-PSDB; AAD36458.
XX
XX Screening animals to determine those likely to produce larger litters
XX and improved meat quality traits involves assaying for the presence of
XX polymorphisms in the AMP activated protein kinase regulatory gamma
XX subunit gene
XX
XX Claim 36; Page 96-97; 109pp; English.
XX
XX The invention relates to a method for screening animals to determine
XX those more likely to produce large litters and improved meat quality
XX traits. The method involves assaying for the presence of a genotype
XX in the sample of genetic material obtained from animal. The genotype
XX is characterized by polymorphism(s) in the AMP activated protein
XX kinase regulatory gamma subunit (PRKAG3) gene. The method is used
XX for screening animals e.g., pigs to determine those most likely to
XX exhibit improved meat quality traits and to produce larger litters.
XX The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).
XX
XX Sequence 464 AA;
XX
XX Query Match 63.6%; Score 136; DB 23; Length 464;
XX Best Local Similarity 65.9%; Pred. No. 4.8e-09;
XX Matches 27; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
XX
XX 1 GKRAKALRWTRQKSVEEGPGGEGPGRSRPAAESTGLEA 41
XX I : I I I I I I I I I I I I I I I I I I I I
XX 26 GDQGNKASRWTRQEDVEEGPGGPGHESPGSRPVAESTQGEA 66
XX
XX RESULT 10
XX ABG03666
XX ID AEG03666 standard; Protein; 171 AA.
XX
XX AC ABG03666;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #3657.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX

```

```

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AA567853.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20; SEQ ID No 34025; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 171 AA;
XX
XX Query Match 30.1%; Score 64.5; DB 22; Length 171;
XX Best Local Similarity 37.2%; Pred. No. 2;
XX Matches 16; Conservative 5; Mismatches 17; Indels 5; Gaps 1;
XX
XX 3 RRAKALRWTRQKSVEEGPGGEG-----PRSRPAAESTGLE 40
XX I : I I I I I I I I I I I I I I I I I I I I
XX 119 RVARGPRSTRSAVDAGPPPSASPGFPPXSSQQRPSPEKTGSE 161
XX
XX RESULT 11
XX AAG35967
XX ID AAG35967 standard; Protein; 61 AA.
XX
XX AC AAG35967;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Zea mays protein fragment SEQ ID NO: 44011.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX

```

PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132663.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.

RESULT 12
AAW71562
ID    AAW71562 standard; Protein; 415 AA.
XX
AC
XX
AAW71562;
XX
DT    21-DEC-1998 (first entry)
XX
DE    Human hepatocyte nuclear factor 1 alpha (truncated mutant).
XX
KW    Hepatocyte nuclear factor 1 alpha; HNF-1 alpha; MODY3; human
KW    transcription factor; maturity onset diabetes of the young;
KW    diabetes; NIDDM; diagnosis; therapy.
XX
OS    Homo sapiens.
XX
Key    Location/Qualifiers
FT    Domain
1..32
/note= "dimerisation domain"
FT    Domain
150..280
/note= "DNA binding domain"
FT    Domain
281..415
/note= "truncated transactivation domain"
XX
WO9811254-A1.
PX    PN
PD    19-MAR-1998.
XX
PF    10-SEP-1997; 97WO-US16037.
XX

PR	30-OCT-1996;	96US-0029679.
PR	10-SEP-1996;	96US-0025719.
PR	02-OCT-1996;	96US-0028056.
PA	(ARCH-)	ARCH DEV CORP.
PI	Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S;	
PI	Oda N, Yamagata K;	
XX	WPI; 1998-271667/24.	
DR	N-PSDB; AAV52628.	
DR		
XX		
PT	Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and	
PT	1-beta - useful for detecting susceptibility for non-insulin	
PT	dependent diabetes, especially maturity-onset diabetes of the young	
XX	Disclosure; Page 180-181; 363pp; English.	
XX		
CC	This is the amino acid sequence of a truncated form of human	
CC	hepatocyte nuclear factor 1 alpha (HNF-1 alpha, see also AAW71559)	
CC	that arises from a frameshift mutation at codon Pro-379 of the	
CC	coding region of the HNF-1 alpha gene (see AAV52626). The mutation	
CC	has been detected in patients with the MODY3 (maturity onset	
CC	diabetes of the young) form of diabetes. The invention concerns	
CC	the identification of genes responsible for non-insulin dependent	
CC	diabetes mellitus (NIDDM) for use in diagnostics and therapeutics.	
CC	It demonstrates that the MODY3 locus is the HNF-1 alpha gene, the	
CC	MODY4 locus is the HNF-1 beta gene (see AAV52730) and the MODY1 locus	
CC	is the HNF-4 alpha gene (see AAV52687). Analysis of mutations in	
CC	these HNF genes can be diagnostic for diabetes. The invention also	
CC	contemplates methods of screening for modulators of HNF function	
CC	utilising HNF nucleic acids or polypeptides, the modulators being	
CC	useful for treating diabetes by modulating HNF function in an	
XX	animal.	
XX		
SQ	Sequence 415 AA;	
Query Match	28.7%; Score 61.5; DB 19; Length 415;	
Best Local Similarity	39.5%; Pred.No.13;	
Matches 15;	Conservative 4; Mismatches 18; Indels 1; Gaps	
QY	2 KRRAKALRWTRQKSVEEGPPGGGEGPRSRPAESTGL 39	
	: : : : : : : : : : : : : : : : : :	
Db	271 RRKEAFRLKLMATYSGPPPGPGPGP-ALPAHSPGL 307	
RESULT 13		
AAW71560		
ID	AAW71560 standard; Protein; 630 AA.	
XX	AAW71560;	
XX		
DT	21-DEC-1998 (first entry)	
XX		
DE	Human hepatocyte nuclear factor 1 alpha (R131Q mutant).	
XX		
KW	Hepatocyte nuclear factor 1 alpha; HNF-1 alpha; MODY3; human;	
KW	transcription factor; maturity onset diabetes of the young;	
KW	diabetes; NIDDM; diagnosis; therapy.	
XX		
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FT	Domain 1..32	
FT	/note= "dimerisation domain"	
FT	Domain 150..280	
FT	/note= "DNA binding domain with POU-like and	
FT	homeodomain-like motifs"	
FT	Domain 281..631	
FT	/note= "transactivation domain"	
FT	Misc-difference 322	
FT	/note= "encoded by TNT"	
XX		

PN WO9811254-A1.  
 XX 19-MAR-1998.  
 XX  
 XX 10-SEP-1997; 97WO-US16037.  
 XX  
 XX 30-OCT-1996; 96US-0025679.  
 PR 10-SEP-1996; 96US-0025719.  
 PR 02-OCT-1996; 96US-0028056.  
 XX  
 XX (ARCH-) ARCH DEV CORP.  
 XX  
 XX Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S;  
 PI Oda N, Yamagata K;  
 XX  
 XX WPI: 1998-271667/24.  
 DR N-PSDB; AAV52626.  
 DR  
 XX Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and  
 PT 1-beta - useful for detecting susceptibility for non-insulin  
 PT dependent diabetes, especially maturity-onset diabetes of the young  
 XX  
 XX Disclosure; Page 169-172; 363pp; English.  
 XX  
 XX This is the amino acid sequence of a mutant (R131Q) form of human  
 CC hepatocyte nuclear factor 1 alpha (HNF-1 alpha, see also AAW71559)  
 CC that arises from a missense mutation (G to A) of codon 131 of the  
 CC coding region of the HNF-1 alpha gene (see AAV52626). The mutation  
 CC has been detected in a patient with the MODY3 (maturity onset  
 CC diabetes of the young) form of diabetes. Arg-131 is conserved in  
 CC human, rat, mouse, hamster, chicken, Xenopus and salmon HNF-1 alpha,  
 CC suggesting this residue to be functionally important. The invention  
 CC concerns the identification of genes responsible for non-insulin  
 CC dependent diabetes mellitus (NIDDM) for use in diagnostics and  
 CC therapeutics. It shows that the MODY3 locus is the HNF-1 alpha gene,  
 CC the MODY4 locus is the HNF-1 beta gene (see AAV52730) and the MODY1  
 CC locus is the HNF-4 alpha gene (see AAV52687). Analysis of mutations in  
 CC these HNF genes can be diagnostic for diabetes. The invention also  
 CC contemplates methods of screening for modulators of HNF function  
 CC utilising HNF nucleic acids or polypeptides, the modulators being  
 CC useful for treating diabetes by modulating HNF function in an  
 CC animal.  
 XX  
 XX SQ Sequence 630 AA;  
 Query Match 28.7%; Score 61.5; DB 19; Length 630;  
 Best Local Similarity 39.5%; Pred. No. 20;  
 Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;  
 QY 2 KRRKALRWTRQKSVGEPPQCGEGPRSPAAESTGL 39  
 Db 271 RKKEAFRHKLAWDTYSGPPGPGPGP-ALPAHSSPGL 307  
 :|: :|:| :| | | | | :| | | | |  
 RESULT 14  
 ID AAW71580  
 XX AAW71580 standard; Protein; 631 AA.  
 XX  
 XX AAW71580;  
 AC  
 XX 21-DEC-1998 (first entry)  
 DT  
 XX Human hepatocyte nuclear factor 1 beta.  
 DE  
 XX Hepatocyte nuclear factor 1 beta; HNF-1 beta; MODY4; human;  
 KW transcription factor; maturity onset diabetes of the young;  
 KW diabetes; NIDDM; diagnosis; therapy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9811254-A1.  
 PN 19-MAR-1998.  
 PD

XX 10-SEP-1997; 97WO-US16037.  
 PF  
 XX 30-OCT-1996; 96US-0029679.  
 PR 10-SEP-1996; 96US-0025719.  
 PR 02-OCT-1996; 96US-0028056.  
 XX  
 XX (ARCH-) ARCH DEV CORP.  
 FA  
 XX Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S;  
 PI Oda N, Yamagata K;  
 XX  
 XX WPI: 1998-271667/24.  
 DR N-PSDB; AAV52729.  
 DR  
 XX Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and  
 PT 1-beta - useful for detecting susceptibility for non-insulin  
 PT dependent diabetes, especially maturity-onset diabetes of the young  
 XX  
 XX Claim 94; Page 228-230; 363pp; English.  
 XX  
 XX This is the amino acid sequence of human hepatocyte nuclear factor  
 CC 1 beta (HNF-1 beta), deduced from a partial gene sequence (see  
 CC AAV52729). HNF-1 beta is a homeodomain-containing transcription  
 CC factor. Mutations in HNF are indicative of a propensity to diabetes  
 CC mellitus. The invention concerns the identification of genes  
 CC responsible for non-insulin dependent diabetes mellitus (NIDDM) for  
 CC use in diagnostics and therapeutics. It demonstrates that the  
 CC MODY3 locus is the HNF-1 alpha gene (see AAV52625), the MODY4 locus  
 CC is the HNF-1 beta gene (see also AAV52730) and the MODY1 locus is the  
 CC HNF-4 alpha gene (see AAV52687). Analysis of mutations in these HNF  
 CC genes can be diagnostic for diabetes. The invention also provides  
 CC methods of utilising HNF nucleic acids and polypeptides in screens  
 CC for modulators of HNF function, the modulators being useful for  
 CC treating diabetes by modulating HNF function in an animal.  
 XX  
 XX SQ Sequence 631 AA;  
 Query Match 28.7%; Score 61.5; DB 19; Length 631;  
 Best Local Similarity 39.5%; Pred. No. 20;  
 Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;  
 QY 2 KRRKALRWTRQKSVGEPPQCGEGPRSPAAESTGL 39  
 Db 271 RKKEAFRHKLAWDTYSGPPGPGPGP-ALPAHSSPGL 307  
 :|: :|:| :| | | | | :| | | | |  
 RESULT 15  
 ID AAW71559  
 XX AAW71559 standard; Protein; 631 AA.  
 XX  
 XX AAW71559;  
 AC  
 XX 21-DEC-1998 (first entry)  
 DT  
 XX Human native hepatocyte nuclear factor 1 alpha.  
 DE  
 XX Hepatocyte nuclear factor 1 alpha; HNF-1 alpha; MODY3; human;  
 KW transcription factor; maturity onset diabetes of the young;  
 KW diabetes; NIDDM; diagnosis; therapy.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Location/Qualifiers  
 FH Key  
 FT Domain 1..32  
 FT /note= "dimerisation domain"  
 FT Domain 150..280  
 FT /note= "DNA binding domain with pOU-like and  
 FT homeodomain-like motifs"  
 FT Domain 281..631  
 FT /note= "transactivation domain"  
 FT Misc-difference 322  
 FT /note= "encoded by TNT"



```
XX WO9811254-A1.
XX PD
XX 19-MAR-1998.
XX PF 10-SEP-1997; 97WO-US16037.
XX 30-OCT-1996; 96OS-0029679.
XX PR 10-SEP-1996; 96US-0025719.
XX 02-OCT-1996; 96US-0028056.
XX (ARCH-) ARCH DEV CORP.
XX PA
XX Bell GI, Furuta H, Horikawa Y, Kaisaki PJ, Menzel S;
XX Oda N, Yamagata K;
XX WI; 1998-271667/24.
XX DR N-PSDB; AAV52625.
XX PT Isolated nucleic acid encoding hepatocyte nuclear factor 1-alpha and
XX 1-beta - useful for detecting susceptibility for non-insulin
XX dependent diabetes, especially maturity-onset diabetes of the young
XX Claim 14; Fig 11; 363pp; English.
XX This is the amino acid sequence of human hepatocyte nuclear factor 1
XX alpha (HNF-1 alpha), a homeodomain-containing sequence-specific DNA
XX binding protein. A cDNA sequence (see AAV52624) encoding HNF-1 alpha
XX is provided. Mutations in this transcription factor are indicative
XX of a propensity to diabetes mellitus. The HNF-1 alpha gene is
XX located on human chromosome 12q, which is the location site of the
XX MODY3 (maturity onset diabetes of the young) locus. The invention
XX concerns the identification of genes responsible for non-insulin
XX dependent diabetes mellitus (NIDDM) for use in diagnostics and
XX therapeutics. It demonstrates that the MODY3 locus is the HNF-1
XX alpha gene, the MODY4 locus is the HNF-1 beta gene (see AAV52730) and
XX the MODY1 locus is the HNF-4 alpha gene (see AAV52687). Analysis of
XX mutations in these HNF genes can be diagnostic for diabetes. The
XX invention also contemplates methods of screening for modulators of
XX HNF function utilising HNF nucleic acids or polypeptides, the
XX modulators being useful for treating diabetes by modulating HNF
XX function in an animal.
XX SQ Sequence 631 AA:
Query Match 28.7%; Score 61.5; DB 19; Length 631;
Best Local Similarity 39.5%; Pred. No. 20;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;
QY 2 KRRKALRWTRQKSVGEPPGGGEGPRSRPAESTGI 39
Db :|:| | | | | | | | | | | | | | |
271 RRKEAFRHLAMDYSGPPPGPGP-ALPAHSPGL 307
```

Search completed: June 6, 2003, 11:01:02  
Job time : 8.60596 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:58:29 ; Search time 2.65674 Seconds  
(without alignments)  
1483.588 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GKRRALKWTRQKSVESCE.....PQGGCPSPRPAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	62	29.0	493	2 T05105	hypothetical prote
2	61.5	28.7	494	2 S39607	transcription fact
3	61.5	28.7	542	2 S39608	transcription fact
4	61.5	28.7	628	1 A39262	transcription fact
5	61.5	28.7	628	1 A33333	transcription fact
6	61.5	28.7	631	1 A36749	transcription fact
7	59	27.6	705	2 A35464	synapsin I splice
8	57.5	26.9	565	2 A36349	pseudouridylylate sy
9	57	26.6	149	1 K5MXE	ribosomal protein
10	56	26.2	435	2 H87516	D-alanyl-D-alanine
11	56	26.2	969	2 A75634	McrB-related prote
12	55.5	25.9	673	2 S36336	probable transcrip
13	55.5	25.9	943	2 T34847	probable transcrip
14	55	25.7	691	2 A25704	synapsin I - rat
15	55	25.7	704	2 A30411	synapsin Ia - rat
16	55	25.7	1085	2 F96712	hypothetical prote
17	55	25.7	1980	2 S54307	myosin heavy chain
18	54.5	25.5	858	2 S68227	ubiquitin thiolest
19	54.5	25.5	1240	2 S21086	anion exchange pro
20	54	25.2	171	2 S55461	mucin - human (fif
21	54	25.2	396	2 D69378	3-ketoacyl-CoA thi
22	54	25.2	701	2 S06475	phenylalanine ammo
23	53.5	25.0	1087	2 T31100	probable potassium
24	53	24.8	203	2 S76108	hypothetical prote
25	52.5	24.5	206	2 B4841	antigen (C-termina
26	52.5	24.5	275	1 JQ1113	interleukin-2 rece
27	52.5	24.5	277	2 T51975	C 3.4.25.1 proteas
28	52.5	24.5	303	2 E88990	protein C36C5.13 [
29	52.5	24.5	323	2 S16318	homeotic protein H

30 52.5 24.5 356 1 WJHU2H homeotic protein H  
31 52.5 24.5 660 1 Q0BE3 BHLFI protein - hu  
32 52 24.3 78 2 C95995 hypothetical prote  
33 52 24.3 369 1 C70666 probable membrane-  
34 52 24.3 517 2 T28630 Y4cC protein - Rhi  
35 52 24.3 612 2 S53714 probable dinitrifi  
36 52 24.3 751 2 T46517 hypothetical prote  
37 52 24.3 1234 2 A34911 band 3-related pro  
38 52 24.3 1237 2 A31789 band 3-related pro  
39 52 24.3 1237 2 A56764 band 3-related pro  
40 52 24.3 1280 2 T00365 hypothetical prote  
41 51.5 24.1 200 2 S14277 NADH2 dehydrogenas  
42 51.5 24.1 259 2 T34536 hypothetical prote  
43 51.5 24.1 363 1 OZZOAK circumporozoite p  
44 51.5 24.1 1359 2 T34036 hypothetical prote  
45 51.5 24.1 1530 1 S01393 DNA-directed RNA p

#### ALIGNMENTS

##### RESULT 1

T05105

hypothetical protein F28M20.160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Mar-2002

C:Accession: T05105

R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mea

submitted to the Protein Sequence Database, November 1998

A:Reference number: Z15398

A:Accession: T05105

A:Molecule type: DNA

A:Residues: 1-493 <BEV>

A:Cross-references: EMBL:AL031004

A:Experimental source: Cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 105/1; 146/3; 296/3; 406/3

A>Note: F28M20.160

C:Supertfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 29.0%; Score 62; DB 2; Length 493;

Best Local Similarity 35.6%; Pred. NO. 9.1;

Matches 16; Conservative 7; Mismatches 8; Indels 14; Gaps 2;

QY 3 RRKALRWTF-----RQKSVESGEPPGQGGPSRPAESTG 38

||| ||| ||| ||| ||| ||| ||| ||| |||

DB 238 KRARVQRWQESRSKHDVREKIAEGEP-----SRNRNRSNKSTG 277

##### RESULT 2

S39607

transcription factor HNF-1A, splice form C - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 29-Aug-1997

C:Accession: S39607; S39752

R:Bach, I.; Yaniv, M.

EMBO J. 12, 4229-4242, 1993

A:Title: More potent transcriptional activators or a transdominant inhibitor of the

A:Reference number: S39607; MUID:94038905; PMID:7900999

A:Accession: S39607

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-494 <BAC>

A:Cross-references: EMBL:X71347

R:Bach, I.

submitted to the EMBL Data Library, April 1993

A:Reference number: S39752

A:Accession: S39752

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 118-494 <BA2>

A:Cross-references: EMBL:X71347

[illegible]

```
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 199-279 <BAC>
A>Note: only a part of the nucleic acid sequence is shown
C:Genetics:
A:Gene: Hnf-1
A:Introns: 238/2
A>Note: the list of introns is incomplete
C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
C:Function:
A:Description: transcription activator required for the expression of a number of liver-
A>Note: also expressed in other tissues (kidney, stomach, intestine), where it may play
C:Superfamily: transcription factor HNF-1; homeobox homology
C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
F:1-33/Region: dimerization
F:222-277/Domain: homeobox homology <HOX>

Query Match      28.7%; Score 61.5; DB 1; Length 628;
Best Local Similarity 39.5%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVEEGPPGCGEGRSRPAESTGL 39
   : : : : : : : : : : : : : : : : : :
Db 271 RRKEAFRKLMDTYNGPPGPGPGP-ALPAHSSPGL 307

RESULT 6
A36749
A:Description: transcription factor HNF-1A - human
N:Alternate names: hepatic nuclear factor 1-alpha; transcription factor 1, hepatic; tran
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36749
R:Bach, I.; Galcheva-Gargova, Z.; Mattei, M.G.; Simon-Chazottes, D.; Guenet, J.L.; Cereg
Genomics 8, 155-164, 1990
A:Title: Cloning of human hepatic nuclear factor 1 (HNF1) and chromosomal localization o
A:Reference number: A36749; MUID:91184801; PMID:1707031
A:Accession: A36749
A:Molecule type: mRNA
A:Residues: 1-631 <BAC>
A:Cross-references: GB:M57732; NID:g184264; PIDN:AAA88077.1; PID:g184265
C:Genetics:
A:Gene: GDB:TCF1; HNF1; LFB1
A:Cross-references: GDB:125297; OMIM:142410; OMIM:600496
A:Map position: 12q24.3-12q24.3
C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
C:Function:
A:Description: transcription activator required for the expression of a number of liver-
A>Note: also expressed in some other tissues, where it may play other roles
C:Superfamily: transcription factor HNF-1; homeobox homology
C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
F:1-33/Region: dimerization
F:222-277/Domain: homeobox homology <HOX>

Query Match      28.7%; Score 61.5; DB 1; Length 631;
Best Local Similarity 39.5%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVEEGPPGCGEGRSRPAESTGL 39
   : : : : : : : : : : : : : : : : : :
Db 271 RRKEAFRKLMDTYNGPPGPGPGP-ALPAHSSPGL 307

RESULT 7
A35363
C:Species: Homo sapiens (man)
N:synapsin I splice form a - human
C:Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 01-Dec-2000
C:Accession: A35363; B35363; A35805
R:Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A:Title: The structure of the human synapsin I gene and protein.
A:Reference number: A35363; MUID:90243651; PMID:2110562
```

---

```
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 199-279 <BAC>
A>Note: only a part of the nucleic acid sequence is shown
C:Genetics:
A:Gene: Hnf-1
A:Introns: 238/2
A>Note: the list of introns is incomplete
C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
C:Function:
A:Description: transcription activator required for the expression of a number of liver-
A>Note: also expressed in other tissues (kidney, stomach, intestine), where it may play
C:Superfamily: transcription factor HNF-1; homeobox homology
C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
F:1-33/Region: dimerization
F:222-277/Domain: homeobox homology <HOX>

Query Match      28.7%; Score 61.5; DB 1; Length 628;
Best Local Similarity 39.5%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVEEGPPGCGEGRSRPAESTGL 39
   : : : : : : : : : : : : : : : : : :
Db 271 RRKEAFRKLMDTYNGPPGPGPGP-ALPAHSSPGL 307

RESULT 6
A36749
A:Description: transcription factor HNF-1A - human
N:Alternate names: hepatic nuclear factor 1-alpha; transcription factor 1, hepatic; tran
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36749
R:Bach, I.; Galcheva-Gargova, Z.; Mattei, M.G.; Simon-Chazottes, D.; Guenet, J.L.; Cereg
Genomics 8, 155-164, 1990
A:Title: Cloning of human hepatic nuclear factor 1 (HNF1) and chromosomal localization o
A:Reference number: A36749; MUID:91184801; PMID:1707031
A:Accession: A36749
A:Molecule type: mRNA
A:Residues: 1-631 <BAC>
A:Cross-references: GB:M57732; NID:g184264; PIDN:AAA88077.1; PID:g184265
C:Genetics:
A:Gene: GDB:TCF1; HNF1; LFB1
A:Cross-references: GDB:125297; OMIM:142410; OMIM:600496
A:Map position: 12q24.3-12q24.3
C:Complex: homodimer; can also form heterodimers with, for example, HNF-1B
C:Function:
A:Description: transcription activator required for the expression of a number of liver-
A>Note: also expressed in some other tissues, where it may play other roles
C:Superfamily: transcription factor HNF-1; homeobox homology
C:Keywords: DNA binding; heterodimer; homeobox; homodimer; liver; nucleus; transcription
F:1-33/Region: dimerization
F:222-277/Domain: homeobox homology <HOX>

Query Match      28.7%; Score 61.5; DB 1; Length 631;
Best Local Similarity 39.5%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVEEGPPGCGEGRSRPAESTGL 39
   : : : : : : : : : : : : : : : : : :
Db 271 RRKEAFRKLMDTYNGPPGPGPGP-ALPAHSSPGL 307

RESULT 7
A35363
C:Species: Homo sapiens (man)
N:synapsin I splice form a - human
C:Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 01-Dec-2000
C:Accession: A35363; B35363; A35805
R:Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A:Title: The structure of the human synapsin I gene and protein.
A:Reference number: A35363; MUID:90243651; PMID:2110562
```

---

```
A:Accession: A35363
A:Molecule type: DNA
A:Residues: 1-705 <SUE>
A:Cross-references: GB:M58371; GB:J05431
A:Accession: B35363
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659, 'KASPAQAQP' <S02>
A:Cross-references: GB:M58378; GB:J05431
R:Sauerwald, A.; Hoesche, C.; Oschwald, R.; Kilimann, M.W.
J. Biol. Chem. 265, 14932-14937, 1990
A:Title: The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and CAAAT-
A:Reference number: A35805; MUID:90368667; PMID:2118519
A:Accession: A35805
A:Molecule type: DNA
A:Residues: 1-125 <SAU>
A:Cross-references: GB:M55301; NID:g338655; PIDN:AAA60608.1; PID:g553654; GB:J05630
C:Genetics:
A:Gene: GDB:SYN1
A:Cross-references: GDB:119606; OMIM:313440
A:Map position: Xp11.23-Xp11.23
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match      27.6%; Score 59; DB 2; Length 705;
Best Local Similarity 42.3%; Pred. No. 29;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 13 QKSVEEGPPGCGEGRSRPAESTG 38
   : : : : : : : : : : : : : :
Db 516 QQPASQAAPPPTGCGRQSRPVAGGPG 541

RESULT 8
AF3639
pseudouridylate synthase (EC 4.2.1.70) [imported] - Brucella melitensis (strain 16M
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C:Accession: AF3639
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivan
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; L
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella mel
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3639
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL54281.1; PID:g17985257; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI11039
A:Map position: II
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      26.9%; Score 57.5; DB 2; Length 565;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 14; Conservative 6; Mismatches 12; Indels 3; Gaps 1;

QY 7 ALRWTRQKSVEEGPPGCGE---GPRSRPAESTG 38
   : : : : : : : : : : : : : :
Db 313 AVQGRTRREMEEGELPQEGERRAPREREWISSTG 347

RESULT 9
R5MXE
ribosomal protein L19, eR - Methanococcus vannielii
N:Alternate names: ribosomal protein E
C:Species: Methanococcus vannielii
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 22-Jun-1999
C:Accession: S05622
R:Auer, J.; Spicker, G.; Boeck, A.
J. Mol. Biol. 209, 21-36, 1989
A:Title: Organization and structure of the Methanococcus transcriptional unit homol
```

S ribosomes.

A:Reference number: S05611; MUID:90040717; PMID:2530355

A:Accession: A75622

A:Molecule type: DNA

A:Residues: 1-149 <AUE>

A:Cross-references: EMBL:X16720; NID:g44754; PIDN:CAA34698.1; PID:g44766

C:Superfamily: rat ribosomal protein L19

C:Keywords: protein biosynthesis; ribosome

Query Match 26.6%; Score 57; DB 1; Length 149;  
Best Local Similarity 35.3%; Pred. No. 12;  
Matches 12; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 2 KRRAKALRWTRKSKVEEGEPGCGEGRSRPAAE 35  
|||:: ||:: ||:: ||:: ||:: ||:: ||:  
Db 51 KKOPEGISSARKKEVQEKRRKKRGGPSRRGAK 84  
||:: ||:: ||:: ||:: ||:: ||:: ||:

RESULT 10  
H87516  
D-alanyl-D-alanine carboxypeptidase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87516  
R:Nierman, W.C.; Feldbiyung, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: H87516  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-435 <STO>  
A:Cross-references: GB:AE005673; NID:q13423656; PIDN:AAK241132.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2161

Query Match 26.2%; Score 56; DB 2; Length 435;  
Best Local Similarity 41.2%; Pred. No. 41;  
Matches 14; Conservative 5; Mismatches 11; Indels 4; Gaps 1;

QY 2 KRRAKALRWTRKSKVEEGEPGCGEGRSRPAAE 35  
||||: ||:: ||:: ||:: ||:: ||:: ||:: ||:  
Db 286 KRRSRGERTTIAANIYEDEP-----SGPIERPSAE 315  
||||: ||:: ||:: ||:: ||:: ||:: ||:: ||:

RESULT 11  
A75634  
McR-related protein - Deinococcus radiodurans (strain RJ)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: A75634  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RJ.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: A75634  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-969 <WHI>  
A:Cross-references: GB:AE001826; NID:q6460827; PIDN:AAF12568.1; PID:g6460864; TIGR:DRB01  
A:Experimental source: strain RI  
C:Genetics:  
A:Gene: DRB0143  
A:Map position: megaplasmid  
A:Genome: plasmid  
A>Note: plasmid MP1

Query Match 26.2%; Score 56; DB 2; Length 969;  
Best Local Similarity 36.6%; Pred. No. 87;

A:Accession: A25704  
A:Molecule type: mRNA  
A:Residues: 1-691 <MCC>  
A:Cross-references: GB:X04655; NID:g57181; PIDN:CAA28353.1; PID:g57182  
A:Experimental source: brain  
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 25.7%; Score 55; DB 2; Length 591;  
Best Local Similarity 42.3%; Pred. No. 83;  
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 13 OKSVEGEGPPGQGGPRSRPAAESTG 38  
I:| : | ||| :||| |  
Db 502 QQSASQATPMTQGGQSRKPVAGGPG 527

RESULT 15  
A30411  
synapsin Ia - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Oct-1990 #sequence\_revision 23-Oct-1990 #text\_change 01-Dec-2000  
C:Accession: A30411; B30411  
R:Suedhof, T.C.; Czerwik, A.J.; Kao, H.T.; Takei, K.; Johnston, P.A.; Horiuchi, A.; Kana  
Science 245, 1474-1480, 1989  
A:Title: Synapsins: mosaics of shared and individual domains in a family of synaptic ves  
A:Reference number: A30411; MUID:89388265; PMID:2506642  
A:Accession: A30411  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-704 <SUE>  
A:Cross-references: GB:M27812; NID:g206920; PIDN:AAA42145.1; PID:g206921  
A:Accession: B30411  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-659, 'KA', 662, 'PAQAP' <SU2>  
A:Cross-references: GB:M27924; NID:g206932; PIDN:AAA42148.1; PID:g206933  
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 25.7%; Score 55; DB 2; Length 704;  
Best Local Similarity 42.3%; Pred. No. 85;  
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 13 OKSVEGEGPPGQGGPRSRPAAESTG 38  
I:| : | ||| :||| |  
Db 514 QQSASQATPMTQGGQSRKPVAGGPG 539

Search completed: June 6, 2003, 11:03:36  
Job time : 4.65674 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 1.50788 seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214

Sequence: 1 GRRAKALRWTRQKSVERGE.....PQGGPRSPRAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	210	98.1	464	AAKL_HUMAN	Q9ug19 homo sapien
2	142	66.4	464	AAKL_PIG	Q9mnp4 sus scrofa
3	61.5	28.7	628	HNFA_MOUSE	P22361 mus musculus
4	61.5	28.7	628	HNFA_RAT	P15257 rattus norv
5	61.5	28.7	631	HNFA_HUMAN	P20823 homo sapien
6	61.5	28.7	958	HIG_DROME	Q09101 drosophila
7	59	27.6	275	CABA_HUMAN	P57796 homo sapien
8	59	27.6	705	SYNL_HUMAN	P17600 homo sapien
9	57	26.6	149	RLJ9_METVA	P14024 methanococc
10	55.5	25.9	569	U171_HUMAN	Q12980 homo sapien
11	55.5	25.9	673	Z145_HUMAN	Q05516 homo sapien
12	55	25.7	270	SL3P_HUMAN	Q14493 homo sapien
13	55	25.7	670	SYNL_MOUSE	O88935 mus musculus
14	55	25.7	704	SYNL_RAT	P09951 rattus norv
15	55	25.7	880	GUNA_THERPU	P28221 thermomonos
16	55	25.7	1980	MY9E_RAT	Q63358 rattus norv
17	54.5	25.5	299	CRX_HUMAN	Q43186 homo sapien
18	54.5	25.5	858	UBP5_HUMAN	P45974 homo sapien
19	54.5	25.5	858	UBP5_MOUSE	P56399 mus musculus
20	54	25.2	701	PAL1_ORYSA	P14717 oryza sativ
21	54	25.2	1739	CHD2_HUMAN	O14647 homo sapien
22	53	24.8	461	Y514_HUMAN	O60269 homo sapien
23	53	24.8	830	SREC_HUMAN	Q14162 homo sapien
24	52.5	24.5	275	IL2A_SHEEP	P26698 ovis aries
25	52.5	24.5	277	PS12_ARATH	O23712 arabidopsis
26	52.5	24.5	299	CRX_BOVIN	Q9xsk0 bos taurus
27	52.5	24.5	299	CRX_MOUSE	O54751 mus musculus
28	52.5	24.5	323	HXDB_MOUSE	P22813 mus musculus
29	52.5	24.5	356	HXB2_HUMAN	P14652 homo sapien
30	52.5	24.5	660	YHL1_EBV	P03181 Epstein-Bar
31	52	24.3	369	MODC_MYCTU	P95155 mycobacteri
32	52	24.3	415	SYNL_CANFA	O62732 canis fami
33	52	24.3	517	Y4CC_RHISN	P55385 rhizobium s

34	52	24.3	528	1	PPGA_HUMAN	Q96134 homo sapien
35	52	24.3	612	1	Y525_PSEAE	Q51484 pseudomonas
36	52	24.3	1234	1	B3A2_RAT	P23347 rattus norv
37	52	24.3	1237	1	B3A2_MOUSE	P13808 mus musculus
38	52	24.3	1237	1	B3A2_RABIT	P48746 oryctolagus
39	52	24.3	1241	1	B3A2_HUMAN	P04920 homo sapien
40	52	24.3	2373	1	CCAH_HUMAN	O95180 homo sapien
41	51.5	24.1	180	1	RS5_SYNPE	O24705 synechococc
42	51.5	24.1	200	1	NUJM_NEUCR	P25710 neurospora
43	51.5	24.1	363	1	CSP_PLAKH	P02894 plasmodium
44	51.5	24.1	478	1	GSR2_HUMAN	Q9nmz5 homo sapien
45	51.5	24.1	1359	1	ATRX_CABEL	Q9u7e0 caenorhabdi

## ALIGNMENTS

RESULT 1  
AAKL\_HUMAN  
ID AAKL\_HUMAN STANDARD; PRT; 464 AA.  
AC Q9UG19; Q9NRL1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)  
DE (AMPK gamma3).  
DE PRKAG3 OR AMPK3.  
GN Homo sapiens (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20164049; PubMed=10698692;  
RA Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;  
RT "Characterization of AMP-activated protein kinase gamma-subunit  
isoforms and their role in AMP binding."  
RL Biochem. J. 345:659-669(2000).

RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=20280150; PubMed=10818001;  
RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Theilander M.,  
RA Rogel-Gaillard C., Paul S., Fannuccelli N., Rask L., Ronne H.,  
RA Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,  
RA Andersson L.;  
RT "A mutation in PRKAG3 associated with excess glycogen content in pig  
skeletal muscle."  
RL Science 288:1248-1251(2000).  
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE  
IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.  
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
CC -!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART  
AND PANCREAS.  
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
SUBUNIT FAMILY.  
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ249977; CAB65117.1; ALT\_INIT.  
DR EMBL; AF214519; AAF73987.1; -.

DR Genew: HGNC:9387; PRKAG3.  
 DR MIM: 604976; -.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam: PF00571; CBS; 4.  
 DR SMART: SM00116; CBS; 4.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT DOMAIN 147 201 CBS 1.  
 FT DOMAIN 228 282 CBS 2.  
 FT DOMAIN 303 356 CBS 3.  
 FT DOMAIN 375 428 CBS 4.  
 FT CONFLICT 58 58 T -> A (IN REF. 1).  
 FT CONFLICT 163 164 MQ -> IE (IN REF. 1).  
 FT CONFLICT 398 398 Q -> K (IN REF. 1).  
 FT CONFLICT 461 464 ALGA -> PSPEKI (IN REF. 1).  
 SQ SEQUENCE 464 AA: 51514 MW: 53985C2C77003A63 CRC64;  
 Query Match 98.1%; Score 210; DB 1; Length 464;  
 Best local similarity 97.6%; Pred. No. 1.1e-16;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVGEPPGQGGPRSPRAAESTGLEA 41  
 |||||  
 DB 26 GKRRKALRWTRQKSVGEPPGQGGPRSPRAAESTGLEA 66  
 RESULT 2  
 AAKL\_PIG  
 ID AAKL\_PIG STANDARD; PRT; 464 AA.  
 AC Q9MYP4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)  
 DE (AMPK gamma3).  
 GN PRKAG3  
 OS Sus scrofa (Pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 CC NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT RN(-) GLN-200.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=20280150; PubMed=10818001;  
 RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,  
 RA Rogel-Gaillard C., Paul S., Iannuccielli N., Rask L., Ronne H.,  
 RA Lundstrom K., Reinsch N., Geilun J., Kalm E., Le Roy P., Chardon P.,  
 RA Andersson L.;  
 RT "A mutation in PRKAG3 associated with excess glycogen content in pig  
 skeletal muscle.";  
 RL Science 288:1248-1251(2000).  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE  
 CC IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -!- TISSUE SPECIFICITY: MUSCLE.  
 CC -!- DISEASE: DEFECTS IN PRKAG3 (RN-) ARE ASSOCIATED WITH EXCESS  
 CC GLYCOGEN CONTENT (ABOUT 70%) IN SKELETAL MUSCLE. THIS MUTATION  
 CC ORIGINATED IN THE HAMPSHIRE BREED PIGS AND HAS BENEFICIAL EFFECTS  
 CC ON MEAT CONTENT BUT DETRIMENTAL EFFECTS ON PROCESSING YIELD. THUS,  
 CC THIS MUTATION IS OF CONSIDERABLE ECONOMIC SIGNIFICANCE IN THE PIG  
 CC BREEDING INDUSTRY.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF214521; AAF73989.1; -.  
 DR EMBL: AF214520; AAF73988.1; -.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam: PF00571; CBS; 4.  
 DR SMART: SM00116; CBS; 4.  
 KW Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation.  
 FT DOMAIN 172 226 CBS 1.  
 FT DOMAIN 253 307 CBS 2.  
 FT DOMAIN 328 381 CBS 3.  
 FT DOMAIN 400 453 CBS 4.  
 FT VARIANT 200 200 R -> Q (IN RN-).  
 SQ SEQUENCE 464 AA: 51308 MW: 17638CB12A2BA9DF CRC64;  
 Query Match 66.4%; Score 142; DB 1; Length 464;  
 Best local similarity 68.3%; Pred. No. 5.3e-09;  
 Matches 28; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 GKRRKALRWTRQKSVGEPPGQGGPRSPRAAESTGLEA 41  
 | : |||||  
 DB 26 GDQGNKASRWTRQEDVEEGPPGQGGPRSPRAAESTGLEA 66  
 RESULT 3  
 HNFA\_MOUSE  
 ID HNFA\_MOUSE STANDARD; PRT; 628 AA.  
 AC P22361;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver specific  
 DE transcription factor LF-B1) (LFb1).  
 DE TCF1 OR HNF1A OR HNF-1A OR HNF-1.  
 GN Mus musculus (Mouse).  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91088607; PubMed=2263635;  
 RA Kuo C.J., Conley P.B., Hsieh C.L., Francke U., Crabtree G.R.;  
 RT "Molecular cloning, functional expression, and chromosomal  
 RT localization of mouse hepatocyte nuclear factor 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9838-9842(1990).  
 CC -!- FUNCTION: REQUIRED FOR THE EXPRESSION OF SEVERAL LIVER SPECIFIC  
 CC GENES. BINDS TO THE INVERTED PALINDROME 5'-GTTAATNATTAAC-3'.  
 CC -!- SUBUNIT: BINDS DNA AS A DIMER.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: LIVER.  
 CC -!- SIMILARITY: BELONGS TO THE HNF1 HOMEOBOX FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M57966; AAA37821.1; -.  
 DR PIR: A39262; A39262.  
 DR HSSP: P15257; ILF6.  
 DR MGD: MGI:98504; Tcf1.  
 DR InterPro: IPR001356; Homeobox.  
 DR SMART: SM00389; HOX; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 DR PROSITE: PS00071; HOMEBOX\_2; 1.  
 KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;  
 KW Activator; Trans-acting factor.



```

FT DOMAIN 1 31 DIMERIZATION.
FT DOMAIN 71 80 ASP/GLU-RICH (ACIDIC) (POTENTIAL
FT FT INVOLVEMENT WITH TRANSCRIPTION).
FT FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 197 205 HOMEBOX.
FT DNA_BIND 199 279
FT DOMAIN 238 258 21 AMINO ACID LOOP BETWEEN HELIX 2 AND 3.
SQ SEQUENCE 628 AA: 67237 MW: 737920D1A369B9DD CRC64;

Query Match 28.7%; DB 1; Length 628;
Best Local Similarity 39.5%; Pred. NO. 8.9;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

OY 2 KRKALRWTKQSVGEPPGCGGPRSPRAESTCL 39
DB 271 RKKEAPRKLMADTYNGPPGPGP-ALPAHSSPL 307

RESULT 4
HNFA_RAT
ID HNFA_RAT STANDARD; PRT; 628 AA.
AC P15257;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver specific
DE transcription factor LF-B1) (LFBI).
GN TCF1 OR HNF1A OR HNF-1A OR HNF-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90003224; PubMed=25714119;
RA Frain M., Swart G., Monaci P., Nicotia A., Staempfli S., Frank R.,
RA Cortese R.;
RT "The liver-specific transcription factor LF-B1 contains a highly
RT diverged homeobox DNA binding domain.";
RL Cell 59:145-157(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91016026; PubMed=2216777;
RA Chouard T., Blumenfeld M., Bach I., Vandekerckhove J., Cereghini S.,
RA Yaniv M.;
RT "A distal dimerization domain is essential for DNA-binding by the
RT atypical HNF1 homeobox domain.";
RN [3]
RP Nucleic Acids Res. 18:5853-5863(1990).
RX SEQUENCE OF 166-628 FROM N.A.
RX MEDLINE=90249741; PubMed=1970973;
RA Baumhueter S., Mendel D.B., Conley P.B., Kuo C.J., Turk C.,
RA Graves M.K., Edwards C.A., Courtis G., Crabtree G.R.;
RT "HNF-1 shares three sequence motifs with the POU domain proteins and
RT is identical to LF-B1 and APF.";
RN [4]
RP Genes Dev. 4:372-379(1990).
RX SEQUENCE OF 1-12 FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Tomei L., Piaggio G., Toniatti C., Lazzaro D., de Francesco R.,
RA Pozzi L., Gerstner J., Cortese R.;
RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
RN [5]
RP POSITION OF HOMEBOX.
RX MEDLINE=90106643; PubMed=1967225;
RA Finney M.;
RT "The homeodomain of the transcription factor LF-B1 has a 21 amino
RT acid loop between helix 2 and helix 3.";
RL Cell 60:5-6(1990).
RN [6]
RP STRUCTURE BY NMR OF 1-32.
RX MEDLINE=9105074; PubMed=1988016;
RA Pastore A., de Francesco R., Barbato G., Castiglione Morelli M.A.,

```

```

RA Motta A., Cortese R.;
RT "1H resonance assignment and secondary structure determination of the
RT dimerization domain of transcription factor LFBI.";
RL Biochemistry 30:148-153(1991).
RN [7]
RP STRUCTURE BY NMR OF 195-286.
RX MEDLINE=93259120; PubMed=8491172;
RA Leitling B., de Francesco R., Tomei L., Cortese R., Otting G.,
RA Wuethrich K.;
RT "The three-dimensional NMR-solution structure of the polypeptide
RT fragment 195-286 of the LFBI/HNF1 transcription factor from rat liver
RT comprises a nonclassical homeodomain.";
RL EMBL J. 12:1797-1803(1993).
RN [8]
RP STRUCTURE BY NMR OF 195-286.
RX MEDLINE=97272000; PubMed=9126845;
RA Schott O., Biller M., Leitling B., Wider G., Wuethrich K.;
RT "The NMR solution structure of the non-classical homeodomain from the
RT rat liver LFBI/HNF1 transcription factor.";
RL J. Mol. Biol. 267:673-683(1997).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 195-286.
RX MEDLINE=93259121; PubMed=8491173;
RA Ceska T.A., Lamers M., Monaci P., Nicotia A., Cortese R., Suck D.;
RT "The X-ray structure of an atypical homeodomain present in the rat
RT liver transcription factor LFBI/HNF1 and implications for DNA
RT binding.";
RL EMBL J. 12:1805-1810(1993).
CC -!- FUNCTION: REQUIRED FOR THE EXPRESSION OF SEVERAL LIVER SPECIFIC
CC GENES. BINDS TO THE INVERTED PALINDROME 5'-GTTAATNATTAAC-3'.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- SIMILARITY: BELONGS TO THE HNF1 HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03170; AAA1524.1; -.
CC EMBL; X54423; CAA38295.1; -.
CC EMBL; X67649; CAA47891.1; -.
CC EMBL; X53297; CAA37387.1; ALT_INIT.
CC PIR; A33333; A33333.
CC PIR; A34590; A34590.
CC PIR; S11568; S11568.
CC PIR; S25485; S25485.
CC PDB; 1LFB; 31-OCT-93.
CC PDB; 2LFB; 12-MAR-97.
CC TRANSFAC; T00369; -.
CC InterPro; IPR001356; Homeobox.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW Activator; Trans-acting factor; 3D-structure.
FT DOMAIN 1 31
FT DOMAIN 71 80 ASP/GLU-RICH (ACIDIC) (POTENTIAL
FT INVOLVEMENT WITH TRANSCRIPTION).
FT FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT FT HOMEBOX.
FT DOMAIN 197 205 21 AMINO ACID LOOP BETWEEN HELIX 2 AND 3.
FT DOMAIN 238 258
FT DOMAIN 208 218
FT TURN 219 220
FT TURN 226 241
FT TURN 242 245
FT TURN 248 249
FT TURN 251 254
FT TURN 255 257
FT HELIX

```

```

FT HELIX      261 273
FT TURN      274 275
SQ SEQUENCE  629 AA; 28713 MW;  8D28099308C86A52 CRC64;

Query Match      28.7%; Score 61.5; DB 1; Length 628;
Best Local Similarity 39.5%; Pred. No. 8.9;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRRAKALRWTRQKSVSEGGPPQGGPRSPRAESTGL 39
DB 271 RKKEAFRHKLAMDYNGPPGPGPGP-ALPAHSSPCL 307

RESULT 5
HNFA_HUMAN
ID HNFA_HUMAN STANDARD; PRT; 631 AA.
AC P20823; Q99861;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte nuclear factor 1-alpha (HNF-1A) (Liver specific
DE transcription factor LF-B1) (LFB1) (Transcription factor-1) (TCF-1).
GN TCF1 OR HNF1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-91184801; PubMed-1707031;
RA Bach I., Galcheva-Gargova Z., Mattei M.-G., Simon-Chazottes D.,
RA Guenet J.-L., Cereghini S., Yaniv M.;
RT "Cloning of human hepatic nuclear factor 1 (HNF1) and chromosomal
RT localization of its gene in man and mouse.";
RL Genomics 8:155-164(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-97100943; PubMed-8945470;
RA Yamagata K., Oda N., Kaisaki P.J., Menzel S., Furuta H.,
RA Vaxillaire M., Southam L., Cox R.D., Lathrop G.M., Boriraj V.V.,
RA Chen X., Cox N.J., Oda Y., Yano H., le Beau M.M., Yamada S.,
RA Nishigori H., Takeda J., Fajans S.S., Hattersley A.T., Iwasaki N.,
RA Hansen T., Pedersen O., Polonsky K.S., Turner R.C., Velho G.,
RA Chevre J.-C., Froguel P., Bell G.I.;
RT "Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-
RT onset diabetes of the young (MODY3).";
RL Nature 384:455-458(1996).
RN [4]
RP VARIANTS MODY3 ARG-107; TRP-131; MET-260 AND HIS-272.
RX MEDLINE-97309269; PubMed-916684;
RA Glucksmann M.A., Lehto M., Tayber O., Scotti S., Berkemeier L.,
RA Pulido J.C., Wu Y., Nir W.-J., Fang L., Markel P., Munnelly K.B.,
RA Goranson J., Orho M., Young B.M., Whitacre J.L., McMenimen C.,
RA Wantman J., Tuomi T., Warim J., Forsblom C.M., Carlsson M.,
RA Rosenzweig J., Kennedy G., Duyk G.M., Krolewski A.S., Groop L.C.,
RA Thomas J.D.;
RT "Novel mutations and a mutational hotspot in the MODY3 gene.";
RL Diabetes 46:1081-1086(1997).
RN [5]
RP VARIANTS MODY3 H-12; Q-131; Q-205 AND C-263, AND VARIANT NIDDM D-191.
RX MEDLINE-97431548; PubMed-9287053;
RA Iwasaki N., Oda N., Ogata M., Hara M., Hinokio Y., Oda Y.,
RA Yamagata K., Kanematsu S., Ohgawara H., Omori Y., Bell G.I.;

```

---

```

RT "Mutations in the hepatocyte nuclear factor-1alpha/MODY3 gene in
RT Japanese subjects with early- and late-onset NIDDM.";
RL Diabetes 46:1504-1508(1997).
RN [6]
RP VARIANTS NIDDM MET-254, AND VARIANTS LEU-27 AND ASN-487.
RX MEDLINE-97431550; PubMed-9287055;
RA Yamada S., Nishigori H., Onda H., Takahashi K., Kitano N.,
RA Morikawa A., Takeuchi T., Takeda J.;
RT "Mutations in the hepatocyte nuclear factor-1alpha gene (MODY3) are
RT not a major cause of late-onset NIDDM in Japanese subjects.";
RL Diabetes 46:1512-1513(1997).
RN [7]
RP VARIANTS IDDM HIS-272 AND GLY-583.
RX MEDLINE-97458988; PubMed-9313763;
RA Yamada S., Nishigori H., Onda H., Utsugi T., Yanagawa T., Maruyama T.,
RA Onigata K., Nagashima K., Nagai R., Morikawa A., Takeuchi T.,
RA Takeda J.;
RT "Identification of mutations in the hepatocyte nuclear factor (HNF)-1-
RT alpha gene in Japanese subjects with IDDM.";
RL Diabetes 46:1643-1647(1997).
RN [8]
RP VARIANTS MODY3, AND VARIANT ATYPICAL DIABETES SER-574.
RX MEDLINE-98052398; PubMed-9392505;
RA Boutin P., Chevre J.-C., Hani E.H., Gomis R., Pardini V.C.,
RA Guillausseau P.-J., Vaxillaire M., Velho G., Froguel P.;
RT "An automated fluorescent single-strand conformation polymorphism
RT technique for screening mutations in the hepatocyte nuclear
RT factor-1alpha gene (maturity-onset diabetes of the young).";
RL Diabetes 46:2108-2109(1997).
RN [9]
RP VARIANTS MODY3 GLN-131; GLN-229; GLY-241 AND HIS-272.
RX MEDLINE-97184312; PubMed-9032114;
RA Kaisaki P.J., Menzel S., Lindner T., Oda N., Rjasanowski I., Sahn J.,
RA Melnicke P., Schulze J., Schmechel H., Petzold C., Ledermann H.M.,
RA Sachse G., Boriraj V.V., Menzel R., Kerner W., Turner R.C.,
RA Yamagata K., Bell G.I.;
RT "Mutations in the hepatocyte nuclear factor-1alpha gene in MODY and
RT early-onset NIDDM: evidence for a mutational hotspot in exon 4.";
RL Diabetes 46:528-535(1997).
RN [10]
RP VARIANTS MODY3 THR-129; TRP-131; TRP-159; LEU-519 AND ILE-620.
RX MEDLINE-97230240; PubMed-9075818;
RA Praying T.M., Bulman M.P., Ellard S., Appleton M., Dronsfield M.J.,
RA Mackie A.D., Baird J.D., Kaisaki P.J., Yamagata K., Bell G.I.,
RA Bain S.C., Hattersley A.T.;
RT "Mutations in the hepatocyte nuclear factor-1alpha gene are a common
RT cause of maturity-onset diabetes of the young in the U.K.";
RL Diabetes 46:720-725(1997).
RN [11]
RP VARIANTS MODY3 ASN-128; TYR-143 AND LEU-447.
RX MEDLINE-97230241; PubMed-9075819;
RA Hansen T., Elberg H., Rouard M., Vaxillaire M., Moeller A.M.,
RA Rasmussen S.K., Fridberg M., Urhammer S.A., Holst J.J., Almind K.,
RA Echwald S.M., Hansen L., Bell G.I., Pedersen O.;
RT "Novel MODY3 mutations in the hepatocyte nuclear factor-1alpha gene:
RT evidence for a hyperexcitability of pancreatic beta-cells to
RT intravenous secretagogues in a glucose-tolerant carrier of a P447L
RT mutation.";
RL Diabetes 46:726-730(1997).
RN [12]
RP VARIANTS LEU-27; VAL-98 AND ASN-487.
RX MEDLINE-97278987; PubMed-9133564;
RA Urhammer S.A., Fridberg M., Hansen T., Rasmussen S.K., Moeller A.M.,
RA Clausen J.O., Pedersen O.;
RT "A prevalent amino acid polymorphism at codon 98 in the hepatocyte
RT nuclear factor-1alpha gene is associated with reduced serum C-peptide
RT and insulin responses to an oral glucose challenge.";
RL Diabetes 46:912-916(1997).
RN [13]
RP VARIANTS NIDDM GLN-583, AND VARIANTS LEU-27; VAL-98 AND ASN-487.
RX MEDLINE-97266232; PubMed-9112026;
RA Urhammer S.A., Rasmussen S.K., Kaisaki P.J., Oda N., Yamagata K.,
RA Moeller A.M., Fridberg M., Hansen L., Hansen T., Bell G.I.;

```



DR Pfam: PF00036; efhand; 3.  
DR PRINTS: PD01362; CALEFLAGIN.  
DR ProDom: PD000012; EF-hand; 2.  
DR SMART: SM00054; EFh; 3.  
DR PROSITE: PS00018; EF\_HAND; 3.  
KW Calcium-binding; Repeat.  
FT CA\_BIND 142 153 EF-HAND 1 (POTENTIAL).  
FT DOMAIN 178 189 ANCESTRAL CALCIUM SITE 2 (POTENTIAL).  
FT CA\_BIND 219 230 EF-HAND 3 (POTENTIAL).  
FT CA\_BIND 256 267 EF-HAND 4 (POTENTIAL).  
SQ SEQUENCE 275 AA; 30433 MW; 1AE849C93AD67DB7 CRC64;

Query Match 27.6%; Score 59; DB 1; Length 275;  
Best Local Similarity 29.5%; Pred.No. 7.5; Indels 12; Gaps 1;  
Matches 13; Conservative 7; Mismatches 12;

QY 3 RRAKALRWTRKQSVBEGE-----PFGQEGGPRSRPAA 34  
: : || : : : || || || || |  
Db 43 KKERLGRSKRKTSGGQTGTGPEAGPSNNPPTGEGPAGAPPA 86

RESULT 8  
SYNL\_HUMAN  
ID SYNL\_HUMAN STANDARD; PRT; 705 AA.  
AC P17600; O75825;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Synapsin I (Brain protein 4.1).  
DE Synapsin I (Brain protein 4.1).  
GN SYNL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90243651; PubMed=2110562;  
RT Suedhof T.C.;  
RT "The structure of the human synapsin I gene and protein.";  
RL J. Biol. Chem. 265:7849-7852(1990).  
RN [2]  
RP SEQUENCE OF 1-125 FROM N.A.  
RX MEDLINE=90368667; PubMed=2118519;  
RX Sauerwald A., Hoese C., Oschwald R., Kilmann M.W.;  
RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and  
RT CAAT-less, phylogenetically conserved sequence with cell  
RT type-specific promoter function.";  
RL J. Biol. Chem. 265:14932-14937(1990).  
RN [3]  
RP SEQUENCE OF 1-258 FROM N.A.  
RA Gratham D.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: NEURONAL PHOSPHOPROTEIN THAT COATS SYNAPTIC VESICLES,  
CC BINDS TO THE CYTOSKELETON, AND IS BELIEVED TO FUNCTION IN THE  
CC REGULATION OF NEUROTRANSMITTER RELEASE.  
CC -!- SUBCELLULAR LOCATION: SYNAPSE.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SYNAPSIN IA (SHOWN HERE) AND  
CC SYNAPSIN IB; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- PTM: SUBSTRATE OF AT LEAST FOUR DIFFERENT PROTEIN KINASES. IT IS  
CC PROBABLE THAT PHOSPHORYLATION PLAYS A ROLE IN THE REGULATION OF  
CC SYNAPSIN I IN THE NERVE TERMINAL.  
CC -!- SIMILARITY: BELONGS TO THE SYNAPSIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M58378; AAC41930.1; -.







DR	Pfam; PF02078; Synapsin; 1.
DR	Pfam; PF02750; Synapsin_C; 1.
DR	PRINTS; PR01368; SYNAPSIN.
DR	PROSITE; PS00415; SYNAPSIN_1; 1.
DR	PROSITE; PS00416; SYNAPSIN_2; 1.
KW	Synapse; Phosphorylation; Nerve; Neurone; Repeat; Actin-binding;
KW	Alternative splicing.
FT	DOMAIN 1 28
FT	DOMAIN 29 112
FT	DOMAIN 113 420
FT	
FT	DOMAIN 421 655
FT	DOMAIN 656 704
FT	MOD_RES 9
FT	MOD_RES 566 566
FT	MOD_RES 603 603
FT	VARSPLIC 660 668
FT	VARSPLIC 669 704
FT	SEQUENCE 704 AA; 73988 MW; 65799FEF7CFE18B5 CRC64;
QY	Query Match 25.7%; Score 55; DB 1; Length 704;
BQ	Best Local Similarity 42.3%; Pred. No. 54;
M	Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
D	13 QKSVEEGPPGGQEGPRSPAAESTG 38   :   :     :   :   :   514 QQSAQAQTPTQQGQRSPVAGGPG 539
R	RESULT 15
I	GUN4.THFU
ID	GUN4.THFU STANDARD; PRT; 880 AA.
AC	P26221; Q08167;
DT	01-MAY-1992 (Rel. 22, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Endoglucanase E-4 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase E-4)
DE	(Cellulase E-4) (Cellulase B4).
OS	CeLd.
OS	Thermomonospora fusca.
OC	Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC	Actinomycetales; Streptosporangineae; Nocardiopsaceae; Thermobifida.
OX	NCBI_TaxID=2021;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=YX;
RC	MEDLINE=94028932; PubMed=8215374;
RA	Jung E.D., Lao G., Irwin D., Barr B.K., Benjamin A., Wilson D.B.;
RT	"DNA sequences and expression in Streptomyces lividans of an
RT	exoglucanase gene and an endoglucanase gene from Thermomonospora
RT	fusca.";
RT	Appl. Environ. Microbiol. 59:3032-3043(1993).
RN	[2]
RN	REVISITONS.
RA	Wilson D.B.;
RN	Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RN	PARTIAL SEQUENCE FROM N.A.
RC	STRAIN=YX;
RC	MEDLINE=91258320; PubMed=1904434;
RA	Lao G., Ghangas G.S., Jung E.D., Wilson D.B.;
RT	"DNA sequences of three beta-1,4-endoglucanase genes from
RT	Thermomonospora fusca.";
RT	J. Bacteriol. 173:3397-3407(1991).
RN	[4]
RN	SEQUENCE OF 47-67.
RA	Wilson D.B.;
RT	"Cellulases of Thermomonospora fusca.";
RL	Meth. Enzymol. 160:314-323(1988).
RN	[5]
RN	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 47-651.
RX	MEDLINE=97475222; PubMed=9334746;
RA	Sakon J., Irwin D., Wilson D.B., Karplus P.A.;



KT "Structure and mechanism of endo/exocellulase E4 from Thermomonospora  
RT fusca.";  
RL Nat. Struct. Biol. 4:810-818(1997).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucoosidic  
CC linkages in cellulose.  
CC -1- PATHWAY: Cellulose degradation.  
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL  
CC HYDROLASES).  
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L20093; ABA42155.1; .  
DR EMBL; M73322; AAA27397.1; ALT\_SEQ.  
DR PIR; B42360; B42360.  
DR PDB; 1J54; 17-SEP-97.  
DR PDB; 1TF4; 04-SEP-97.  
DR PDB; 3TF4; 04-SEP-97.  
DR PDB; 4TF4; 04-SEP-97.  
DR InterPro; IPR001919; Bac\_celose-bind.  
DR InterPro; IPR001956; CBD\_3.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR003962; FNIII\_repeat.  
DR InterPro; IPR001701; GH\_9.  
DR InterPro; IPR001230; Prenyl\_site.  
DR Pfam; PF00041; fn3; 1.  
DR Pfam; PF00553; CBM\_2; 1.  
DR Pfam; PF00759; Glyco\_hydro\_9; 1.  
DR Pfam; PF00942; CBM\_3; 1.  
DR PRINTS; PR00014; FNTYPEIII.  
DR SMART; SM00060; FN3; 1.  
DR PROSITE; PS00561; CBD\_BACTERIAL; 1.  
DR PROSITE; PS00592; GLYCOSYL\_HYDROL\_F9\_1; 1.  
DR PROSITE; PS00598; GLYCOSYL\_HYDROL\_F9\_2; 1.  
KW Cellulose degradation; Hydrolase; Glycosidase; Signal; 3D-structure.  
FT SIGNAL 1 46  
FT CHAIN 47 880 ENDOGLUCANASE E-4.  
FT DOMAIN 776 880 CELLULOSE-BINDING.  
FT ACT\_SITE 427 427 BY SIMILARITY.  
FT ACT\_SITE 461 461 BY SIMILARITY.  
FT ACT\_SITE 470 470 BY SIMILARITY.  
SQ SEQUENCE 880 AA; 95202 MW; 5EA9A6ABF45A4D9A CRC64;  
  
Query Match 25.7%; Score 55; DB 1; Length 880;  
Best Local Similarity 83.3%; Pred. NO. 67;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 17 EEGEPPGGEGP 28  
Db 652 EEGEPPGGEGP 663  
|||||  
|||||

Search completed: June 6, 2003, 11:01:30  
Job time : 3.50788 secs

Result No.	Query Score	Query			DB ID	Description
		Match	Length	DB		
1	65	30.4	399	11	Q8VCQ7 mus musculus	
2	62	29.0	493	10	O81782 arabidopsis	
3	61.5	28.7	866	5	Q917E3 drosophila	
4	61.5	28.7	933	5	Q917E4 drosophila	
5	61.5	28.7	958	5	Q9V560 drosophila	
6	60	28.0	136	10	Q42174 arabidopsis	
7	60	28.0	314	10	Q8S835 oryza sativ	
8	60	28.0	486	10	Q9LDJ7 arabidopsis	
9	60	28.0	793	4	Q9H5Z2 homo sapien	
10	59.5	27.8	330	10	Q8S5D8 oryza sativ	
11	59.5	27.8	1400	5	Q9NKS7 leishmania	
12	59	27.6	514	5	Q961A3 drosophila	
13	59	27.6	863	5	Q9VRV3 drosophila	
14	58.5	27.3	207	15	Q8US87 human immun	
15	58	27.1	277	16	Q910D4 streptomyce	
16	58	27.1	303	16	Q98GW6 rhizobium l	

OS *Arabidopsis thaliana* (Mouse-ear)

OS Aral



```

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003834; AAG22296.1; -.
DR HSSP: P10998; 1VVD.
DR FlyBase: FBgn0010114; hig.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00084; sushi; 5.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 5.
DR SMART: SM00409; IG; 1.
SQ SEQUENCE 933 AA; 104178 MW; 90CB696854C6C9FA CRC64;

Query Match 28.7%; Score 61.5; DB 5; Length 933;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 14; Conservative 8; Mismatches 11; Indels 5; Gaps 2;

Qy 3 RRAKALRWTRQKSV-EGE-----PPGQEGPRSRPAAE 35
Db 328 QDRSPWRRRRSTFKGDLQGLPPESGIGPEPEPLAD 365

RESULT 5
Q9V560 PRELIMINARY; PRT; 958 AA.
AC Q9V560;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HG protein.
GN HG OR CG2040.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Cocayne J.D.,

```

```

RA Ananatzides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003834; AAF58960.1; -.
DR HSSP: P10998; 1VVD.
DR FlyBase: FBgn0010114; hig.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00084; sushi; 5.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 5.
DR SMART: SM00409; IG; 1.
SQ SEQUENCE 958 AA; 107027 MW; 7B2A08477672B674 CRC64;

Query Match 28.7%; Score 61.5; DB 5; Length 958;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 14; Conservative 8; Mismatches 11; Indels 5; Gaps 2;

Qy 3 RRAKALRWTRQKSV-EGE-----PPGQEGPRSRPAAE 35
Db 328 QDRSPWRRRRSTFKGDLQGLPPESGIGPEPEPLAD 365

RESULT 6
Q42174 PRELIMINARY; PRT; 136 AA.
AC Q42174;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Globulin (fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```



```

RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026410; BAB13475.1; -.
DR InterPro; IPR005112; dDENN.
DR InterPro; IPR001194; DENN.
DR InterPro; IPR005113; uDENN.
DR Pfam; PF03455; dDENN; 1.
DR Pfam; PF02141; DENN; 1.
DR Pfam; PF03456; uDENN; 1.
SQ SEQUENCE 793 AA; 86230 MW; 3F37ACA836F24AF4 CRC64;

Query Match      28.0%; Score 60; DB 4; Length 793;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 16; Conservative 4; Mismatches 8; Indels 4; Gaps 2;

QY 1 GKRRKALRWTRQKSVVEG--EPFGQGEGRPS 30
   || | || |:: |||| |||| | |
Db 504 GKNR--PLRPSRRRQLEGTSEPPGAGTPPLS 533

RESULT 10
Q85SD8 PRELIMINARY; PRT; 330 AA.
AC Q85SD8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 37.7 kDa protein.
GN OSJNB0047B19.26.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nascimento L., Zutavern T., Ballia V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB0047B19, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC113339; AAM08880.1; -.
KW Hypothetical protein.
SQ SEQUENCE 330 AA; 37674 MW; 01AD287515E7DD43 CRC64;

Query Match      27.8%; Score 59.5; DB 10; Length 330;
Best Local Similarity 37.2%; Pred. No. 13;
Matches 16; Conservative 5; Mismatches 13; Indels 9; Gaps 2;

QY 2 KRRKALRWTRQKSVVEGEPGQGEGRPS---RPAAESTGLEA 41
   :|| |||| | | | | | | | | | | | | | | | |
Db 38 RRRFAALRMWR-----GGPPVTARGPQATARQPLATMRGSOA 74

RESULT 11
Q9NKS7 PRELIMINARY; PRT; 1400 AA.
AC Q9NKS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE L509.6.
GN L509.6.
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Myler P.J.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005836; AAF35994.1; -.
SQ SEQUENCE 1400 AA; 149577 MW; 0CDCF4E925BEAEF4 CRC64;

Query Match      27.8%; Score 59.5; DB 5; Length 1400;
Best Local Similarity 36.4%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 11; Indels 17; Gaps 1;

QY 12 RQKSVVEGEP-----PGQGEGRPSRPAAESTG 38
   || || || | | | | | | | | | | | |
Db 432 RQPEEEEEPRRAFLKDDAVLAEGVQPADGEAAAGRAPASSTG 475

RESULT 12
Q961A3 PRELIMINARY; PRT; 514 AA.
AC Q961A3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE LD26701p.
GN CG10274.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051740; AAK93164.1; -.
DR FlyBase; FBgn0035690; CG10274.
DR InterPro; IPR000822; Znf.C2H2.
DR InterPro; IPR000130; Zn_Mp1ptidse.
DR Pfam; PF00096; zf-C2H2; 9.
DR PROSITE; PS00028; ZINC_FINGER_C2H2.1; UNKNOWN_7.
DR PROSITE; PS00157; ZINC_FINGER_C2H2.2; 9.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 514 AA; 59029 MW; 42DC53A6A4EC5F49 CRC64;

Query Match      27.6%; Score 59; DB 5; Length 514;
Best Local Similarity 35.9%; Pred. No. 24;
Matches 14; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 2 KRRKALRWTRQKSVVEGEPGQGEGRPSRPAAESTGLE 40
   || |:: || | | | | | | | | | | |
Db 183 KRVLRAIYRQVRQVEMGETPGEQFNPPAPPVEGISIE 221

RESULT 13
Q9VRV3 PRELIMINARY; PRT; 863 AA.
AC Q9VRV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG10274 protein.
GN CG10274.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aqbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bouchan M.A., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Daveyport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fodor C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., McIntosh A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidden-Klamas I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs K.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Sclentre 287:2185-2195(2000).
DR EMBL: AE003563; AAF50679.1; -.
DR FlyBase: Fbm0035690; CG10274.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000130; Znf_MTPeptase.
DR Pfam: PF00096; Znf_C2H2; 12.
DR SMART: SM00455; Znf_C2H2; 12.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 12.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 863 AA; 98670 MW; 6A47464249835106 CRC64;

Query Match 27.6%; Score 59; DH 5; Length 863;
Best Local Similarity 35.9%; Prod. No. 42;
Matches 14; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

Qy 2 KRKALRWTRKQSVSEGEPPGQGGPRSKPAANSTGLE 40
Db 532 KRLVLAETIYRQGVEMGFTPGKQFNDNPAFTVKGTSIE 570

RESULT 14
Q8US87 PRELIMINARY; PRT; 207 AA.
AC Q8US87;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gag polyprotein (fragment).
GN Gag.
OS Human immunodeficiency virus type 1.

Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=99CMLB33;
MEDLINE=21663212; PubMed=11804555;
RA Tebit D.M., Zekeng L., Kaptue L., Salminen M., Krausslich H.G.,
RA Herchenroder O.;
RT "Genotypic and Phenotypic Analysis of HIV Type 1 Primary Isolates from
Western Cameroon."
RL AIDS Res. Hum. Retroviruses 18:39-48(2002).
DR EMBL: AF352594; AAL68724.1; -.
DR InterPro: IPR000721; Gag_p24.
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR Pfam: PF00607; Gag_p24; 1.
KW Polyprotein.
FT NON_TER 1 207
FT NON_TER 207 207
SQ SEQUENCE 207 AA; 23082 MW; D99E86CA6421EE21 CRC64;

Query Match 27.3%; Score 58.5; DB 15; Length 207;
Best Local Similarity 42.4%; Prod. No. 11;
Matches 14; Conservative 3; Mismatches 15; Indels 1; Gaps 1;

Qy 6 KALRWTRKQSVSEGEPPGQGGPRSKPAANST 37
Db 175 EAAEWDLHPVQAGPHPPGQLRDRGRDIAGTT 207
: | | | | | | | | | | | | | | | | | | | |
: | | | | | | | | | | | | | | | | | | | |

RESULT 15
Q9L0D4 PRELIMINARY; PRT; 277 AA.
AC Q9L0D4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 30S ribosomal protein S3.
GN RFSC OR SC04708 OR SCD31.33.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2);
MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood B.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

```

RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)";  
 RL Nature 417:141-147(2002).  
 DR EMBL; ALJ161803; CAB62076.1; .  
 DR InterPro; IPR004087; KH\_dom.  
 DR InterPro; IPR004044; KH\_TYPE\_2.  
 DR InterPro; IPR001351; Ribosomal\_S3.  
 DR Pfam; PF00013; KH-domain; 1.  
 DR Pfam; PF00189; Ribosomal\_S3\_C; 1.  
 DR Pfam; PF00417; Ribosomal\_S3\_N; 1.  
 DR SMART; SM00322; KH; 1.  
 DR TIGRFAMs; TIGR01009; rpsc\_bact; 1.  
 DR PROSITE; PS50823; KH\_TYPE\_2; 1.  
 DR PROSITE; PS00548; RIBOSOMAL\_S3; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 277 AA; 30273 MW; 5831536E64010135 CRC64;

Query Match 27.1%; Score 58; DB 16; Length 277;  
 Best Local Similarity 41.5%; Pred. No. 17;  
 Matches 17; Conservative 5; Mismatches 15; Indels 4; Gaps 2;

QY 1 GKRRKALRWTRQKSVGEPPGQGGPRSRPAALSTGLEA 41  
 |:|| | :|| | | : | | : |||| |  
 Db 241 GERRG---RKPOQUSAPAAEAP-KAEAPAAAAAESTGTTEA 277

Search completed: June 6, 2003, 11:02:51  
 Job time : 7.31349 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 2.44133 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91  
Perfect score: 214  
Sequence: 1 GKRRKALWTRKQSVERGE.....PGQGGPRSRPAESTGLEA 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61.5	28.7	415	4	US-08-927-219-8
2	61.5	28.7	630	4	US-08-927-219-2
3	61.5	28.7	630	4	US-08-927-219-4
4	61.5	28.7	631	4	US-08-927-219-127
5	60	28.0	187	4	US-09-199-637A-287
6	59	27.6	696	3	US-08-906-865-4
7	59	27.6	696	4	US-09-129-668-4
8	52	24.3	2353	4	US-08-984-709A-50
9	51	23.8	499	2	US-08-820-170A-40
10	51	23.8	499	3	US-09-055-699-40
11	51	23.8	499	4	US-09-273-565-40
12	51	23.8	499	4	US-09-565-538-40
13	51	23.8	499	4	US-09-661-468-40
14	51	23.8	503	4	US-09-562-737-62
15	51	23.8	1251	5	PCT-US95-02251-3
16	51	23.8	1252	1	US-08-199-780-3
17	51	23.8	1252	2	US-08-316-650-3
18	51	23.8	1253	3	US-08-479-722B-4
19	50.5	23.6	223	4	US-09-206-676C-1
20	50.5	23.6	226	4	US-09-206-676C-2
21	50.5	23.6	1497	1	US-08-623-679-7
22	50.5	23.6	1497	3	US-08-933-774-7
23	50.5	23.6	1497	4	US-09-181-030-7
24	50.5	23.6	1497	4	US-09-534-242-7
25	50.5	23.6	1497	4	US-09-454-854-7
26	50.5	23.6	1497	4	US-09-164-671-7
27	50.5	23.6	1533	1	US-08-623-679-9

28	50.5	23.6	1533	3	US-08-933-774-9
29	50.5	23.6	1533	4	US-09-181-030-9
30	50.5	23.6	1533	4	US-09-534-242-9
31	50.5	23.6	1533	4	US-09-454-854-9
32	50.5	23.6	1533	4	US-09-164-671-9
33	50	23.4	410	1	US-07-945-283-4
34	50	23.4	449	1	US-09-041-075A-21
35	50	23.4	4928	4	US-09-036-987A-5
36	50	23.4	4928	4	US-09-370-700-5
37	49.5	23.1	106	2	US-08-785-065-5
38	49.5	23.1	106	4	US-09-151-412-5
39	49.5	23.1	233	2	US-08-458-568A-4
40	49.5	23.1	448	4	US-09-310-463-18
41	49.5	23.1	448	4	US-08-842-248A-18
42	49	22.9	595	4	US-09-370-838-187
43	49	22.9	882	4	US-09-413-814-78
44	49	22.9	1780	1	US-08-769-309A-5
45	49	22.9	1780	3	US-08-994-570-5

ALIGNMENTS

RESULT 1  
US-08-927-219-8  
; Sequence 8, Application US/08927219  
; Patent No. 6187533  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Graeme I.  
; APPLICANT: Yamagata, Kazuya  
; APPLICANT: Oda, Naohisha  
; APPLICANT: Kaisaki, Pamela J.  
; APPLICANT: Furuta, Hiroto  
; APPLICANT: Horikawa, Yukio  
; APPLICANT: Menzel, Stephen  
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY  
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,219  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,679  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/028,056  
; FILING DATE: 02-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/025,719  
; FILING DATE: 10-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mark B.  
; REGISTRATION NUMBER: 37,259  
; REFERENCE/DOCKET NUMBER: ARCD:272  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 415 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-927-219-8

Query Match 28.7%; Score 61.5; DB 4; Length 415;  
Best Local Similarity 39.5%; Pred. No. 2.2;  
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVVEGEPGOGGPRSRPAESTGL 39  
DB 271 RRKEAFRKLAMDTYSGPPGPGP-ALPAHSSPGL 307

RESULT 2

US-08-927-219-2  
; Sequence 2, Application US/08927219  
; Patent No. 6187533  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Graeme I.  
; APPLICANT: Yamagata, Kazuya  
; APPLICANT: Oda, Naohisha  
; APPLICANT: Kaisaki, Pamela J.  
; APPLICANT: Furuta, Hiroto  
; APPLICANT: Horikawa, Yukio  
; APPLICANT: Menzel, Stephen  
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY  
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,219  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,679  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/028,056  
; FILING DATE: 02-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/025,719  
; FILING DATE: 10-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mark B.  
; REGISTRATION NUMBER: 37,259  
; REFERENCE/DOCKET NUMBER: ARCD:272  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 630 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-927-219-2

Query Match 28.7%; Score 61.5; DB 4; Length 630;  
Best Local Similarity 39.5%; Pred. No. 3.6;  
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;  
QY 2 KRAKALRWTRQKSVVEGEPGOGGPRSRPAESTGL 39  
DB 271 RRKEAFRKLAMDTYSGPPGPGP-ALPAHSSPGL 307

RESULT 3

US-08-927-219-4  
; Sequence 4, Application US/08927219  
; Patent No. 6187533  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Graeme I.  
; APPLICANT: Yamagata, Kazuya  
; APPLICANT: Oda, Naohisha  
; APPLICANT: Kaisaki, Pamela J.  
; APPLICANT: Furuta, Hiroto  
; APPLICANT: Horikawa, Yukio  
; APPLICANT: Menzel, Stephen  
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY  
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,219  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,679  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/028,056  
; FILING DATE: 02-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/025,719  
; FILING DATE: 10-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mark B.  
; REGISTRATION NUMBER: 37,259  
; REFERENCE/DOCKET NUMBER: ARCD:272  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 630 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-927-219-4

Query Match 28.7%; Score 61.5; DB 4; Length 630;  
Best Local Similarity 39.5%; Pred. No. 3.6;  
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRAKALRWTRQKSVVEGEPGOGGPRSRPAESTGL 39  
DB 271 RRKEAFRKLAMDTYSGPPGPGP-ALPAHSSPGL 307

RESULT 4

```

US-08-927-219-127
: Sequence 127, Application US/08927219
: Patent No. 6187533
: GENERAL INFORMATION:
: APPLICANT: Bell, Graeme I.
: APPLICANT: Yamagata, Kazuya
: APPLICANT: Oda, Naohisha
: APPLICANT: Kaisaki, Pamela J.
: APPLICANT: Furuta, Hiroto
: APPLICANT: Menzel, Stephen
: APPLICANT: Horikawa, Yukio
: APPLICANT: Menzel, Stephen
: TITLE OF INVENTION: MUTATIONS IN THE DIAPYTES SUSCEPTIBILITY
: TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA
: TITLE OF INVENTION: AND HNF-4ALPHA
: NUMBER OF SEQUENCES: 147
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/927,219
: FILING DATE: Concurrently Herewith
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/029,679
: FILING DATE: 30-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/028,056
: FILING DATE: 02-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/025,719
: FILING DATE: 10-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Wilson, Mark B.
: REGISTRATION NUMBER: 37,259
: REFERENCE/DOCKET NUMBER: ARCD:272
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 512/418-3000
: TELEFAX: 512/474-7577
: INFORMATION FOR SEQ ID NO: 127:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 631 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: US-08-927-219-127

Query Match 28.7%; Score 61.5; DB 4; Length 631;
Best Local Similarity 39.5%; Pred. No. 3.6;
Matches 15; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 KRRAKLRWTRKSVEEGPPGGGGRSRPAAESTGL 39
: : : : : : : : : : : : : : : : : :
Db 271 RRKEEAFRHKLANDTYSGPPPGPGPGP-ALPAHSSPGL 307

RESULT 5
US-09-199-637A-287
: Sequence 287, Application US/09199637A
: Patent No. 6355411
: GENERAL INFORMATION:
: APPLICANT: Ausubel, Frederick
: APPLICANT: Goodman, Howard M.
: APPLICANT: Rahme, Laurence G.
: APPLICANT: Mahajan-Miklos, Shalina

```



```
RESULT 10
US-09-055-699-40
; Sequence 40, Application US/09055699
; Patent No. 6005088
; GENERAL INFORMATION:
; APPLICANT: TSUTOMU, FUJIWARA
; APPLICANT: TAKESHI, WATANABE
; APPLICANT: MASATO, HORIE
; APPLICANT: TOYOMASA, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,699
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/820,170
; APPLICATION FILING DATE:
; FILING DATE:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-055-699-40

Query Match 23.8%; Score 51; DB 3; Length 499;
Best Local Similarity 38.7%; Pred. No. 66;
Matches 12; Conservative 1; Mismatches 16; Indels 2; Gaps 1;

QY 10 WTRQKSVEEGPPGCGEGRSPRAAESTGLE 40
| | | | | | | | | | | | | | | | | | | |
Db 21 WEHSKEVSEAEPPGGSSGDSGPP--EESQGE 49

RESULT 11
US-09-273-565-40
; Sequence 40, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997

Query Match 23.8%; Score 51; DB 3; Length 499;
Best Local Similarity 38.7%; Pred. No. 66;
Matches 12; Conservative 1; Mismatches 16; Indels 2; Gaps 1;

QY 10 WTRQKSVEEGPPGCGEGRSPRAAESTGLE 40
| | | | | | | | | | | | | | | | | | | |
Db 21 WEHSKEVSEAEPPGGSSGDSGPP--EESQGE 49

RESULT 12
US-09-565-538-40
; Sequence 40, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/565,538
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/273,565
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 40
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-565-538-40

Query Match 23.8%; Score 51; DB 4; Length 499;
Best Local Similarity 38.7%; Pred. No. 66;
Matches 12; Conservative 1; Mismatches 16; Indels 2; Gaps 1;

QY 10 WTRQKSVEEGPPGCGEGRSPRAAESTGLE 40
| | | | | | | | | | | | | | | | | | | |
Db 21 WEHSKEVSEAEPPGGSSGDSGPP--EESQGE 49

RESULT 13
US-09-661-468-40
; Sequence 40, Application US/09661468
; Patent No. 6376189
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/661,468
; CURRENT FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 11:02:59 ; Search time 3.37478 Seconds  
(without alignments)  
1254.259 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_51\_91

Perfect score: 214  
Sequence: 1 GKRRAKALRWTRQKSVEEGE.....PGQGGPSRPAESTGLEA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214	100.0	489	10	US-09-826-581-6
2	60	28.0	187	9	US-09-975-719-287
3	60	28.0	270	10	US-09-729-674-132
4	59	27.6	696	9	US-10-122-805-4
5	56.5	26.4	529	12	US-10-042-417-44
6	56	26.2	447	10	US-09-888-615-109
7	55.5	25.9	107	9	US-09-738-626-6443
8	55	25.7	240	9	US-10-042-894A-2
9	54	25.2	1243	9	US-10-196-935A-4
10	53.5	25.0	150	9	US-09-764-891-3210
11	53.5	25.0	150	10	US-09-908-711-123
12	53.5	25.0	217	9	US-10-062-548-113
13	53	24.8	111	10	US-09-864-761-47499
14	53	24.8	205	9	US-10-091-458-38
15	53	24.8	830	9	US-09-870-759-134
16	53	24.8	830	9	US-09-842-758-57
17	52.5	24.5	436	10	US-09-925-300-1640
18	52	24.3	549	10	US-09-764-864-1131
19	52	24.3	830	9	US-09-870-759-140

20	51.5	24.1	478	9	US-09-924-340-108	Sequence 108, App
21	51.5	24.1	478	9	US-09-992-600A-108	Sequence 108, App
22	51.5	24.1	478	9	US-09-746-783-184	Sequence 184, App
23	51.5	24.1	478	9	US-10-000-489-108	Sequence 108, App
24	51.5	24.1	478	9	US-10-000-986-108	Sequence 108, App
25	51	23.8	56	10	US-09-864-761-41214	Sequence 41214, A
26	51	23.8	247	9	US-09-738-626-5773	Sequence 5773, App
27	51	23.8	347	10	US-09-789-561-184	Sequence 184, App
28	51	23.8	459	9	US-10-153-668-86	Sequence 86, Appl
29	51	23.8	499	10	US-09-976-165-40	Sequence 40, Appl
30	51	23.8	503	9	US-10-211-962-62	Sequence 62, Appl
31	50.5	23.6	480	9	US-09-796-753-158	Sequence 158, App
32	50.5	23.6	756	10	US-09-946-175-2	Sequence 2, Appli
33	50.5	23.6	761	10	US-09-946-175-3	Sequence 3, Appli
34	50.5	23.6	792	9	US-10-059-585-22	Sequence 22, Appl
35	50.5	23.6	934	9	US-09-796-753-156	Sequence 156, App
36	50.5	23.6	985	9	US-09-978-295A-211	Sequence 211, App
37	50.5	23.6	985	9	US-09-978-697-211	Sequence 211, App
38	50.5	23.6	985	9	US-09-978-192A-211	Sequence 211, App
39	50.5	23.6	985	9	US-09-999-832A-211	Sequence 211, App
40	50.5	23.6	985	9	US-09-978-189-211	Sequence 211, App
41	50.5	23.6	985	9	US-09-978-608A-211	Sequence 211, App
42	50.5	23.6	985	9	US-09-978-191A-211	Sequence 211, App
43	50.5	23.6	985	9	US-09-978-403A-211	Sequence 211, App
44	50.5	23.6	985	9	US-09-978-564A-211	Sequence 211, App
45	50.5	23.6	985	9	US-09-978-585A-211	Sequence 211, App

ALIGNMENTS

RESULT 1

US-09-826-581-6  
; Sequence 6, Application US/09826581  
; Patent No. US20020142310A1  
; GENERAL INFORMATION:  
; APPLICANT: Andersson, Leif  
; APPLICANT: Luthman, L. Holger  
; APPLICANT: Marklund, Stefan  
; TITLE OF INVENTION: VARIANTS OF THE HUMAN AMP-ACTIVATED PROTEIN KINASE GAMMA 3 SI  
; FILE REFERENCE: 11145-007001  
; CURRENT APPLICATION NUMBER: US/09/826,581  
; CURRENT FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 60/195,665  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-581-6

Query Match 100.0%; Score 214; DB 10; Length 489;

Best local similarity 100.0%; Pred. No. 4.1e-17;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKRRAKALRWTRQKSVEEGPPGQGGPSRPAESTGLEA 41

DB 51 GKRRAKALRWTRQKSVEEGPPGQGGPSRPAESTGLEA 91

RESULT 2

US-09-975-719-287  
; Sequence 287, Application US/09975719  
; Publication No. US20030022349A1  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Rahme, Laurence G.  
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
; FILE REFERENCE: 00786/361003  
; CURRENT APPLICATION NUMBER: US/09/975,719

```
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-287
```

```
Query Match      28.0%; Score 60; DB 9; Length 187;
Best Local Similarity 34.9%; Pred. No. 11;
Matches 15; Conservative 6; Mismatches 16; Indels 6; Gaps 1;
```

```
QY 1 GKRRAKALRWTRQ-----KSVEEGPPQGGGPRSRPAEST 37
      |:| | | | | | | | | | | | | | | | | | | | |
DB 47 GRRANGARWTRLRPRGRSLADAPAPCAAASRARPRASST 89
```

## RESULT 3

```
US-09-729-674-132
; Sequence 132, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallic, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fecthel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-132
```

```
Query Match      28.0%; Score 60; DB 10; Length 270;
Best Local Similarity 36.6%; Pred. No. 16;
Matches 15; Conservative 8; Mismatches 8; Indels 10; Gaps 2;
```

```
QY 3 RRAKALRWTRQKSVEEGPPQGG-----EGP---RSRPA 33
      | | | | | | | | | | | | | | | | | | | | |
DB 219 RDARYQASRKQEQAGQPPHGRGDASSALCQGPVPVGRPA 259
```

## RESULT 4

```
US-10-122-805-4
; Sequence 4, Application US/10122805
; Publication No. US20030082645A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE AND USES
```

```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-202 CIP
; CURRENT APPLICATION NUMBER: US/10/122,805
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/129,668
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/906,865
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-805-4
```

```
Query Match      27.6%; Score 59; DB 9; Length 696;
Best Local Similarity 42.3%; Pred. No. 55;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 13 QKSVEEGPPQGGGPRSRPAESTG 38
      | | | | | | | | | | | | | | | | | | | | |
DB 516 QQPASQAAPPTQGGGQSRPVAGGPG 541
```

## RESULT 5

```
US-10-042-417-44
; Sequence 44, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-44
```

```
Query Match      26.4%; Score 56.5; DB 12; Length 529;
Best Local Similarity 32.7%; Pred. No. 81;
Matches 16; Conservative 4; Mismatches 18; Indels 11; Gaps 1;
```

```
QY 1 GKRRAKALRWTRQK-----SVEEGPPQGGGPRSRPAESTG 38
      | | | | | | | | | | | | | | | | | | | | |
DB 8 GRGREKRARGARRKRGQGREARAAADGGGSGPGAAGARTPRPEAEG 56
```

## RESULT 6

```
US-09-888-615-109
; Sequence 109, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: RHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
```



```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-109

Query Match      26.2%; Score 56; DB 10; Length 447;
Best Local Similarity 55.0%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 19 GEPPGQGGPRSRPAESTG 38
DB 177 GEAPGLGAGPAPSPMSSSTG 196

RESULT 7
US-09-738-626-6443
; Sequence 6443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6443
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6443

Query Match      25.9%; Score 55.5; DB 9; Length 107;
Best Local Similarity 36.7%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 10 WTRQKSV-----EEGPPGQGGPRSRP 32
DB 61 WSRTEVERSPGLPEDGRPPCAAPAPTASP 90

RESULT 8
US-10-042-894A-2
; Sequence 2, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-109

Query Match      26.2%; Score 56; DB 10; Length 447;
Best Local Similarity 55.0%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 19 GEPPGQGGPRSRPAESTG 38
DB 177 GEAPGLGAGPAPSPMSSSTG 196

RESULT 7
US-09-738-626-6443
; Sequence 6443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6443
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6443

Query Match      25.9%; Score 55.5; DB 9; Length 107;
Best Local Similarity 36.7%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 10 WTRQKSV-----EEGPPGQGGPRSRP 32
DB 61 WSRTEVERSPGLPEDGRPPCAAPAPTASP 90

RESULT 8
US-10-042-894A-2
; Sequence 2, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-109

Query Match      26.2%; Score 56; DB 10; Length 447;
Best Local Similarity 55.0%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 19 GEPPGQGGPRSRPAESTG 38
DB 177 GEAPGLGAGPAPSPMSSSTG 196

RESULT 7
US-09-738-626-6443
; Sequence 6443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6443
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6443

Query Match      25.9%; Score 55.5; DB 9; Length 107;
Best Local Similarity 36.7%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 10 WTRQKSV-----EEGPPGQGGPRSRP 32
DB 61 WSRTEVERSPGLPEDGRPPCAAPAPTASP 90

RESULT 8
US-10-042-894A-2
; Sequence 2, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-109

Query Match      26.2%; Score 56; DB 10; Length 447;
Best Local Similarity 55.0%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 19 GEPPGQGGPRSRPAESTG 38
DB 177 GEAPGLGAGPAPSPMSSSTG 196

RESULT 7
US-09-738-626-6443
; Sequence 6443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6443
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6443

Query Match      25.9%; Score 55.5; DB 9; Length 107;
Best Local Similarity 36.7%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 10 WTRQKSV-----EEGPPGQGGPRSRP 32
DB 61 WSRTEVERSPGLPEDGRPPCAAPAPTASP 90

RESULT 8
US-10-042-894A-2
; Sequence 2, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-109

Query Match      26.2%; Score 56; DB 10; Length 447;
Best Local Similarity 55.0%; Pred. No. 78;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 19 GEPPGQGGPRSRPAESTG 38
DB 177 GEAPGLGAGPAPSPMSSSTG 196

RESULT 7
US-09-738-626-6443
; Sequence 6443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

```



```
; FILE REFERENCE: P2024P1
; CURRENT APPLICATION NUMBER: US/10/062,548
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/369,247
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/074,118
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,157
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,137
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,341
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074,141
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (131)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-10-062-548-113

Query Match      25.0%; Score 53.5; DB 9; Length 217;
Best Local Similarity 48.0%; Pred. No. 72;
Matches 12; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY      14 KSVEEGEPGQGGGPRSRPAAEETG 38
      :::: || || || || || || ||
Db      76 RTMDGTGPPAVGE-PRSGPSAGSAG 99

RESULT 13
US-09-864-761-47489
; Sequence 47489, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
```

```
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006655
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47489
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008962.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 14
; OTHER INFORMATION: EST_HUMAN HIT: AW602053.1, EVALUE 7.00e-25
; OTHER INFORMATION: SWISSPROT HIT: Q16696, EVALUE 3.00e-27
US-09-864-761-47489

Query Match      24.8%; Score 53; DB 10; Length 111;
Best Local Similarity 41.7%; Pred. No. 41;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY      9 RWRQKSVGEPPGQGGGPRSRP 32
      : | : | | | | | | | |
Db      65 QWEELQSAEGNQAGSGEGRHPR 88

RESULT 14
US-10-091-458-38
; Sequence 38, Application US/10091458
; Publication No. US20030068627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT206C1
; CURRENT APPLICATION NUMBER: US/10/091,458
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,900
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
```

;  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
;  
;  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17

Search completed: June 6, 2003, 11:16:42  
Job time : 4.37478 secs

```

; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match      24.8%; Score 53; DB 9; Length 205;
Best Local Similarity 32.2%; Pred. No. 78;
Matches 19; Conservative 2; Mismatches 12; Indels 26; Gaps 2;

QY      9  RWTRKQSVFEGEPGQG-----EG----PRSRPAAESTGLEA 41
Db      145  RLTLQAQREGGEVAGQTAHAGRRRVEAANVDELVEGEWLYPRPGPASEPTKLRA 203

RESULT 15
US-09-870-759-134
; Sequence 134, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-134

Query Match      24.8%; Score 53; DB 9; Length 830;
Best Local Similarity 44.0%; Pred. No. 3.3e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY      14  KSVFEGPPGQGGPRSRPAAESTG 38
Db      630  REAESTGPDEAEAPESFPAAASPG 654
```



PR 27-APR-2000; 2000US-0560875.  
XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Drmanac RT;  
PI WPI: 2001-457740/49.  
XX N-PSDB; ABA08485.  
DR Human proteins and DNA encoding sequences useful for preventing,  
XX treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -  
XX Claim 20; Page 159-160; 1963pp; English.  
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis; cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention.  
XX Sequence 181 AA;  
SQ Query Match 100.0%; Score 213; DB 22; Length 181;  
Best Local Similarity 100.0%; Pred. No. 1.6e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
DB 80 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 120  
|||||  
RESULT 2  
AAE00221  
ID AAE00221 standard; Protein: 305 AA.  
XX AC AAE00221;  
XX DT 13-JUN-2001 (first entry)  
XX DE Human AMPK gamma subunit muscle-specific isoform, PRKAG3.

XX Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy.  
XX Homo sapiens.  
OS W0200120003-A2.  
PN 22-MAR-2001.  
XX 11-SEP-2000; 2000WO-EP09896.  
XX 10-SEP-1999; 99EP-0402236.  
PR 18-MAY-2000; 2000EP-0401388.  
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX WPI: 2001-244810/25.  
DR N-PSDB; AAD03296.  
XX New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX Claim 4; Page 55-57; 71pp; English.  
XX The present sequence is human adenosine monophosphate  
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
CC PRKAG3. Mutation in prkag3 results in an altered regulation of  
CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
CC useful as therapeutic for treating carbohydrate metabolism disorders such  
CC as diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3 or a  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
XX Sequence 305 AA;  
SQ Query Match 100.0%; Score 213; DB 22; Length 305;  
Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
DB 136 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 176  
|||||  
RESULT 3  
AAE00328  
ID AAE00328 standard; Protein: 305 AA.  
XX AC AAE00328;  
XX

DT 13-JUN-2001 (first entry)  
 XX Human Prkag3 R41Q mutant.  
 XX Human: gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiatic; gene therapy; mutant; mutein;  
 XX variant.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 PH Misc-difference 41  
 FT /note= "Wild-type Arg substituted with Gln"  
 XX  
 XX WO200120003-A2.  
 XX  
 XX 22-MAR-2001.  
 XX  
 XX 11-SEP-2000; 2000WO-EP09896.  
 XX  
 XX 10-SEP-1999; 99EP-0402236.  
 XX  
 XX 18-MAY-2000; 2000EP-0401388.  
 XX  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX WPI: 2001-244810/25.  
 XX  
 XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 XX Disclosure; Page -: 71pp; English.  
 XX  
 XX The present sequence is a R41Q mutant of human muscle-specific isoform  
 CC of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
 CC (AMPK) Prkag3. This mutant sequence results in increased glycogen  
 CC content in human skeletal muscle.  
 CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the human Prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
 CC of sequence listing (AAE00221).  
 XX  
 XX Sequence 305 AA;  
 SQ  
 Query Match 100.0%; Score 213; DB 22; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 41  
 DE ||||||||||||||||||||||||||||||||||||  
 XX  
 DB 136 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 176  
 XX  
 RESULT 4  
 AAE00329  
 ID AAE00329 standard; Protein; 305 AA.  
 XX  
 XX AAE00329;  
 XX  
 XX 13-JUN-2001 (first entry)  
 DT  
 XX  
 XX Human Prkag3 V40I mutant.  
 DE  
 XX  
 XX Human: gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiatic; gene therapy; mutant; mutein;  
 KW variant.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 PH Misc-difference 40  
 FT /note= "Wild-type Val substituted with Ile"  
 FT  
 XX  
 XX WO200120003-A2.  
 XX  
 XX 22-MAR-2001.  
 XX  
 XX 11-SEP-2000; 2000WO-EP09896.  
 XX  
 XX 10-SEP-1999; 99EP-0402236.  
 XX  
 XX 18-MAY-2000; 2000EP-0401388.  
 XX  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA (ANDE/) ANDERSSON L.  
 PA (LOOF/) LOOFT C.  
 PA (KALM/) KALM E.  
 XX  
 XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX WPI: 2001-244810/25.  
 XX  
 XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders  
 PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX  
 XX Disclosure; Page -: 71pp; English.  
 XX  
 XX The present sequence is a V40I mutant of human muscle-specific isoform  
 CC of gamma subunit of adenosine monophosphate (AMP)-activated kinase  
 CC (AMPK) Prkag3. This mutant sequence results in decreased glycogen  
 CC content in human skeletal muscle.  
 CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
 CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
 CC therapeutic for treating carbohydrate metabolism disorders such as  
 CC diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the human Prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
 CC of sequence listing (AAE00221).  
 XX  
 XX Sequence 305 AA;  
 SQ  
 Query Match 100.0%; Score 213; DB 22; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



CC encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.  
 CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the human prkag3 sequence SEQ.ID.NO.4 shown in page 57-58  
 CC of sequence listing (AAE00221).

XX Sequence 305 AA;  
 Query Match 100.0%; Score 213; DB 22; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 DB 136 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 176

RESULT 5  
 AAE00223  
 ID AAE00223 standard; Protein; 464 AA.  
 XX AC AAE00223;  
 XX DT 13-JUN-2001 (first entry)  
 XX DE Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3.  
 KW Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiac; gene therapy.

XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH 172..225  
 FT Domain /label= CBS  
 FT Misc-difference 200 /note= "Cystathione beta synthase domain"  
 FT /note= "RN- mutation site"  
 FT Domain 253..307  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 329..382  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"  
 FT Domain 400..453  
 FT /label= CBS  
 FT /note= "Cystathione beta synthase domain"

XX WO200120003-A2.  
 XX 22-MAR-2001.  
 XX 11-SEP-2000; 2000WO-EP09896.  
 XX 10-SEP-1999; 99EP-0402236.  
 XX 18-MAY-2000; 2000EP-0401388.  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 XX (ANDE/) ANDERSSON L.  
 XX (LOOF/) LOOFT C.  
 XX (KALM/) KALM E.  
 XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
 PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
 XX WPI: 2001-244810/25.  
 DR N-PSDB; AAD03320.  
 XX New variants of the gamma subunit of vertebrate adenosine  
 PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and  
 PT myopathy -  
 XX Claim 5; Fig 3; 71pp; English.  
 PS The present sequence is human adenosine monophosphate  
 XX (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
 CC complete PRKAG3. Mutation in prkag3 results in an altered regulation of  
 CC carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is  
 CC useful as therapeutic for treating carbohydrate metabolism disorders such  
 CC as diabetes, obesity, and disorders associated with muscle metabolism  
 CC such as myopathy and cardiovascular diseases, to modulate AMPK  
 CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
 CC and its functionally altered mutants are useful for the diagnostic  
 CC evaluation, genetic testing and prognosis of a metabolic disorder,  
 CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.  
 CC transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.

XX Sequence 464 AA;

Query Match 100.0%; Score 213; DB 22; Length 464;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 DB 295 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 6  
 AAB47679  
 ID AAB47679 standard; Protein; 489 AA.  
 XX AC AAB47679;  
 XX DT 21-JAN-2002 (first entry)  
 XX PRKAG3.  
 XX Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;  
 KW metabolic disease; diabetes; obesity; substitution; ss.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 71 /note= "Possible variation point P71A"  
 FT Misc-difference 340 /note= "Possible variation point R340W"  
 FT WO200177305-A2.  
 XX 18-OCT-2001.  
 XX 06-APR-2001; 2001WO-SE00765.  
 XX 07-APR-2000; 2000US-195665P.  
 XX (AREX-) AREXIS AB.

XX Andersson L, Luthman H, Marklund S;  
 PI WPI: 2001-657170/75.  
 DR N-PSDB; RAH43685.  
 XX

PT New variants of human AMP-activated protein kinase gamma3 subunit  
PT associated with a metabolic disease e.g. diabetes or obesity and method  
PT for determining a risk estimate of diseases in subject by detecting the  
PT variant -  
XX  
PS Disclosure; Fig 5; 25pp; English.  
XX  
CC This sequence is encoded by the full length cDNA encoding the human  
CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting  
CC the presence of the PRKAG3 DNA, or a variant, is useful in determining  
CC a risk estimate of a metabolic disease, such as diabetes or obesity,  
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon  
CC 3 variation may be a substitution of a G for a C at nucleotide 320,  
CC resulting in the amino acid substitution P71A; in exon 4 variation may  
CC be a substitution of a T for a C at nucleotide 550; and in exon 10  
CC variation may be a substitution of a T for a C at nucleotide 1037,  
CC resulting in the amino acid substitution R340W. There may also be  
CC nucleotide variation in intron 6. The numbering of these  
CC variations is based on the full length cDNA, rather than on  
CC position 1 of the open reading frame.  
XX  
XX Sequence 489 AA;  
XX  
XX Query Match 100.0%; Score 213; DB 22; Length 489;  
XX Best Local Similarity 100.0%; Pred. No. 4.8e-23;  
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 41  
DB 320 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 360  
  
RESULT 7  
AAE00220  
ID AAE00220 standard; Protein; 305 AA.  
AC AAE00220;  
XX  
XX 13-JUN-2001 (first entry)  
XX  
DE Pig AMPK gamma subunit muscle-specific isoform, PRKAG3.  
XX  
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;  
KW chromosome 15.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FH Domain 13..66  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 94..148  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 170..223  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
FT Domain 241..294  
FT /label= CBS  
FT /note= "Cystathione beta synthase domain"  
XX  
XX WC200120003-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 11-SEP-2000; 2000WO-EP09896.  
XX  
XX 10-SEP-1999; 99EP-0402336.  
XX  
XX 18-MAY-2000; 2000EP-0401388.  
XX

PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
PA (ANDE/) ANDERSSON L.  
PA (LOOF/) LOOFT C.  
PA (KALM/) KALM E.  
XX  
PI Andersson L, Loof C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;  
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;  
XX  
DR WPI: 2001-244810/25.  
XX N-PSDB: AAD03295.  
XX  
PT New variants of the gamma subunit of vertebrate adenosine  
PT monophosphate-activated kinase for diagnosis or treatment of disorders  
PT associated with energy metabolism such as diabetes, obesity, and  
PT myopathy -  
XX  
PS Claim 4; Fig 2; 7lpp; English.  
XX  
CC The present amino acid sequence is pig adenosine monophosphate  
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,  
CC PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.  
CC Mutation in Prkag3 results in an altered regulation of carbohydrate  
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as  
CC therapeutic for treating carbohydrate metabolism disorders such as  
CC diabetes, obesity, and disorders associated with muscle metabolism  
CC such as myopathy and cardiovascular diseases, to modulate AMPK  
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence  
CC and its functionally altered mutants are useful for the diagnostic  
CC evaluation, genetic testing and prognosis of a metabolic disorder,  
CC preferably a carbohydrate metabolism disorder. Primers that can detect  
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
CC from the expression of a functionally altered allele of PRKAG3.  
CC Transgenic animal and host cell transformed with PRKAG3 or a  
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
CC screening compounds able to modulate AMPK activity. Nucleic acid  
CC encoding PRKAG3 is useful for detecting mutations in a prkag3 gene, or  
CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
CC of PRKAG3 and is useful in gene therapy.  
XX  
XX Sequence 305 AA;  
XX  
XX Query Match 98.6%; Score 210; DB 22; Length 305;  
XX Best Local Similarity 97.6%; Pred. No. 7.8e-23;  
XX Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTRDLA 41  
DB 136 HILTHKRLKFLHFGTLPRPSFLYRTIQDLGIGTRDLA 176  
  
RESULT 8  
AAE00225  
ID AAE00225 standard; Protein; 305 AA.  
XX  
AC AAE00225;  
XX  
XX 13-JUN-2001 (first entry)  
XX  
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).  
XX  
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; mutein;  
KW variant.  
XX  
XX Sus scrofa.  
XX  
XX Key Location/Qualifiers  
XX Domain 13..66  
XX /label= CBS  
XX /note= "Cystathione beta synthase domain"

```

FT Misc-difference 41 /note= "Wild-type Arg substituted with Gln"
FT Domain 94..148 /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223 /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294 /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
XX WO200120003-A2.
XX
XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-EP09896.
XX
XX 10-SEP-1999; 99EP-0402236.
XX 18-MAY-2000; 2000EP-0401388.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFT C.
XX (KALM/) KALM E.
XX
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
XX WPI; 2001-244810/25.
XX
XX New variants of the gamma subunit of vertebrate adenosine
XX monophosphate-activated kinase for diagnosis or treatment of disorders
XX associated with energy metabolism such as diabetes, obesity, and
XX myopathy -
XX
XX Claim 9; Page -; 71pp; English.
XX
XX The present sequence is a R41Q mutant of muscle-specific isoform of
XX gamma subunit of adenosine monophosphate (AMP)-activated kinase
XX (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in
XX increased glycogen content in pig skeletal muscle.
XX Mutation in Prkag3 results in an altered regulation of carbohydrate
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX therapeutic for treating carbohydrate metabolism disorders such as
XX diabetes, obesity, and disorders associated with muscle metabolism
XX such as myopathy and cardiovascular diseases, to modulate AMPK
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX and its functionally altered mutants are useful for the diagnostic
XX evaluation, genetic testing and prognosis of a metabolic disorder,
XX preferably a carbohydrate metabolism disorder. Primers that can detect
XX a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
XX useful for detecting a dysfunction of carbohydrate metabolism resulting
XX from the expression of a functionally altered allele of PRKAG3.
XX Transgenic animal and host cell transformed with PRKAG3 or a
XX heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
XX screening compounds able to modulate AMPK activity. Nucleic acid
XX encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
XX in a sequence encoding the first cystathione beta synthase (CBS) domain
XX of PRKAG3 and is useful in gene therapy.
XX Note: The present sequence is not shown in the specification, but
XX is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).
XX
SQ Sequence 305 AA;

Query Match 98.6%; Score 210; DB 22; Length 305;
Best Local Similarity 97.6%; Pred. No. 7.8e-23;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGICTFRDLA 41
Db 136 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGICTFRDLA 176

```

```

RESULT 9
AAE00226
ID AAE00226 standard; Protein; 305 AA.
XX
AC AAE00226;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa AMPK gamma chain isoform Prkag3 mutant (R41Q).
XX
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; mutant; muten;
KW variant.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT Domain 13..66
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Misc-difference 40
FT /note= "Wild-type Val substituted with Ile"
FT Domain 94..148
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 170..223
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
FT Domain 241..294
FT /label= CBS
FT /note= "Cystathione beta synthase domain"
XX
XX WO200120003-A2.
XX
XX 22-MAR-2001.
XX
XX 11-SEP-2000; 2000WO-EP09896.
XX
XX 10-SEP-1999; 99EP-0402236.
XX 18-MAY-2000; 2000EP-0401388.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (ANDE/) ANDERSSON L.
XX (LOOF/) LOOFT C.
XX (KALM/) KALM E.
XX
XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
XX Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
XX WPI; 2001-244810/25.
XX
XX New variants of the gamma subunit of vertebrate adenosine
XX monophosphate-activated kinase for diagnosis or treatment of disorders
XX associated with energy metabolism such as diabetes, obesity, and
XX myopathy -
XX
XX Claim 9; Page -; 71pp; English.
XX
XX The present sequence is a V40I mutant of muscle-specific isoform of
XX gamma subunit of adenosine monophosphate (AMP)-activated kinase
XX (AMPK) Prkag3 from Sus scrofa. This mutant sequence results in
XX decreased glycogen content in pig skeletal muscle.
XX Mutation in Prkag3 results in an altered regulation of carbohydrate
XX metabolism, particularly in skeletal muscle. PRKAG3 is useful as
XX therapeutic for treating carbohydrate metabolism disorders such as
XX diabetes, obesity, and disorders associated with muscle metabolism
XX such as myopathy and cardiovascular diseases, to modulate AMPK
XX activity, and for restoring a normal AMPK function. PRKAG3 sequence
XX and its functionally altered mutants are useful for the diagnostic
XX evaluation, genetic testing and prognosis of a metabolic disorder,
XX

```

CC preferably a carbohydrate metabolism disorder. Primers that can detect  
 CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are  
 CC useful for detecting a dysfunction of carbohydrate metabolism resulting  
 CC from the expression of a functionally altered allele of PRKAG3.

CC Transgenic animal and host cell transformed with PRKAG3 or a  
 CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for  
 CC screening compounds able to modulate AMPK activity. Nucleic acid  
 CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or  
 CC in a sequence encoding the first cystathione beta synthase (CBS) domain  
 CC of PRKAG3 and is useful in gene therapy.

CC Note: The present sequence is not shown in the specification, but  
 CC is derived from the porcine Prkag3 sequence shown in Fig 2 (AAE00220).

XX Sequence 305 AA;

Query Match 98.6%; Score 210; DB 22; Length 305;  
 Best Local Similarity 97.6%; Pred. No. 7.8e-23;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKLFHIFGSLLPSPFLYRTIQDLGIGTRDLA 41  
 |||||  
 Db 136 HILTHKRLKLFHIFGTLPPSPFLYRTIQDLGIGTRDLA 176

RESULT 10

ID AAE00222 standard; Protein; 464 AA.

AC AAE00222;

DT 13-JUN-2001 (first entry)

XX Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3.

XX Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;  
 KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;  
 KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;  
 KW cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;  
 KW chromosome 15.

OS Sus scrofa.

FH Key Location/Qualifiers  
 FT Domain 172..225

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Misc-difference 200

FT /note= "RN- mutation site"

FT Domain 253..307

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Domain 329..382

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

FT Domain 400..453

FT /label= CBS

FT /note= "Cystathione beta synthase domain"

XX WO200120003-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-EP09896.

XX 10-SEP-1999; 99EP-0402236.

XX 18-MAY-2000; 2000EP-0401388.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

PA (ANDE/) ANDERSSON L.

PA (LOOF/) LOOFT C.

PA (KALM/) KALM E.

XX Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;

PI Iannuccielli N, Gellin J, Le Roy P, Chardon P;

XX WPI: 2001-244810/25.

DR N-PSDB; AAD03319.

XX New variants of the gamma subunit of vertebrate adenosine

PT monophosphate-activated kinase for diagnosis or treatment of disorders

PT associated with energy metabolism such as diabetes, obesity, and

PT myopathy -

XX

PS Claim 5; Fig 3; 71pp; English.

XX

XX The present sequence is pig adenosine monophosphate

CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,

CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome

CC 15. Mutation in prkag3 results in an altered regulation of carbohydrate

CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as

CC therapeutic for treating carbohydrate metabolism disorders such as

CC diabetes, obesity, and disorders associated with muscle metabolism

CC such as myopathy and cardiovascular diseases, to modulate AMPK

CC activity, and for restoring a normal AMPK function. PRKAG3 sequence

CC and its functionally altered mutants are useful for the diagnostic

CC evaluation, genetic testing and prognosis of a metabolic disorder,

CC preferably a carbohydrate metabolism disorder. Primers that can detect

CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are

CC useful for detecting a dysfunction of carbohydrate metabolism resulting

CC from the expression of a functionally altered allele of PRKAG3.

CC Transgenic animal and host cell transformed with PRKAG3 or a

CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for

CC screening compounds able to modulate AMPK activity. Nucleic acid

CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or

CC in a sequence encoding the first cystathione beta synthase (CBS) domain

CC of PRKAG3 and is useful in gene therapy.

XX

SQ Sequence 464 AA;

Query Match 98.6%; Score 210; DB 22; Length 464;

Best Local Similarity 97.6%; Pred. No. 1.3e-22;

Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKLFHIFGSLLPSPFLYRTIQDLGIGTRDLA 41

|||||

Db 295 HILTHKRLKLFHIFGTLPPSPFLYRTIQDLGIGTRDLA 335

RESULT 11

AAE22984

ID AAE22984 standard; Protein; 464 AA.

XX AAE22984;

XX 09-AUG-2002 (first entry)

XX Pig wild-type PRKAG3 protein.

XX

KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;

KW screening; meat quality; single nucleotide polymorphism; SNP; pig.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT Misc-difference 30

FT /note= "Wild type Asn is replaced with Thr during

FT single nucleotide polymorphism (SNP)"

FT

FT Misc-difference 52

FT /note= "Wild type Gly is replaced with Ser during

FT single nucleotide polymorphism (SNP)"

FT

FT Misc-difference 199

FT /note= "Wild type Val is replaced with Ile during

FT single nucleotide polymorphism (SNP)"

FT

FT Misc-difference 200

FT /note= "Wild type Arg is replaced with Gln during

FT single nucleotide polymorphism (SNP)"

FT



XX WPI: 2002-393850/42.  
 DR N-PSDB; AAD36458.  
 XX  
 PT Screening animals to determine those likely to produce larger litters  
 PT and improved meat quality traits involves assaying for the presence of  
 PT polymorphisms in the AMP activated protein kinase regulatory gamma  
 PT subunit gene -  
 XX  
 PS Claim 36; Page 96-97; 109pp; English.  
 XX  
 CC The invention relates to a method for screening animals to determine  
 CC those more likely to produce large litters and improved meat quality  
 CC traits. The method involves assaying for the presence of a genotype  
 CC in the sample of genetic material obtained from animal. The genotype  
 CC is characterised by polymorphism(s) in the AMP activated protein  
 CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
 CC for screening animals e.g., pigs to determine those most likely to  
 CC exhibit improved meat quality traits and to produce larger litters.  
 CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-52).  
 XX  
 SQ Sequence 464 AA;  
 Query Match 98.6%; Score 210; DB 23; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 1.3e-22;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 DB 295 HILTHKRLKFLHFGTLLPRPSFLYRTIQDLGIGTFRDLA 335  
 RESULT 14  
 AAEE22987  
 ID AAEE22987 standard; Protein; 464 AA.  
 XX  
 AC AAEE22987;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig PRKAG3 polymorphic variant (PRKAG3-199).  
 XX  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 199  
 FT /note= "Wild type Val is substituted with Ile due  
 FT to single nucleotide polymorphism (SNP)"  
 XX  
 PN W0200220850-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001WO-US28283.  
 XX  
 PR 08-SEP-2000; 2000US-231045P.  
 PR 08-JAN-2001; 2001US-260239P.  
 PR 18-JUN-2001; 2001US-299111P.  
 XX  
 PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 XX  
 PI Rothschild MF, Ciobanu DC, Malek M, Plastow G;  
 XX  
 DR WPI: 2002-393850/42.  
 DR N-PSDB; AAD36459.  
 XX  
 CC Screening animals to determine those likely to produce larger litters  
 PT and improved meat quality traits involves assaying for the presence of  
 PT polymorphisms in the AMP activated protein kinase regulatory gamma

PT subunit gene -  
 XX  
 PS Disclosure; Page 100-102; 109pp; English.  
 XX  
 CC The invention relates to a method for screening animals to determine  
 CC those more likely to produce large litters and improved meat quality  
 CC traits. The method involves assaying for the presence of a genotype  
 CC in the sample of genetic material obtained from animal. The genotype  
 CC is characterised by polymorphism(s) in the AMP activated protein  
 CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
 CC for screening animals e.g., pigs to determine those most likely to  
 CC exhibit improved meat quality traits and to produce larger litters.  
 CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-199).  
 XX  
 SQ Sequence 464 AA;  
 Query Match 98.6%; Score 210; DB 23; Length 464;  
 Best Local Similarity 97.6%; Pred. No. 1.3e-22;  
 Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
 |||||  
 DB 295 HILTHKRLKFLHFGTLLPRPSFLYRTIQDLGIGTFRDLA 335  
 RESULT 15  
 AAEE22988  
 ID AAEE22988 standard; Protein; 464 AA.  
 XX  
 AC AAEE22988;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Pig PRKAG3 polymorphic variant (PRKAG3-200).  
 XX  
 KW AMP activated protein kinase regulatory gamma subunit; PRKAG3 gene;  
 KW screening; meat quality; single nucleotide polymorphism; SNP; pig;  
 KW variant.  
 XX  
 OS Sus scrofa.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 200  
 FT /note= "Wild type Arg is substituted with Gln due  
 FT to single nucleotide polymorphism (SNP)"  
 XX  
 PN W0200220850-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001WO-US28283.  
 XX  
 PR 08-SEP-2000; 2000US-231045P.  
 PR 08-JAN-2001; 2001US-260239P.  
 PR 18-JUN-2001; 2001US-299111P.  
 XX  
 PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 XX  
 PI Rothschild MF, Ciobanu DC, Malek M, Plastow G;  
 XX  
 DR WPI: 2002-393850/42.  
 DR N-PSDB; AAD36460.  
 XX  
 CC Screening animals to determine those likely to produce larger litters  
 PT and improved meat quality traits involves assaying for the presence of  
 PT polymorphisms in the AMP activated protein kinase regulatory gamma  
 PT subunit gene -  
 XX  
 PS Disclosure; Page 105-107; 109pp; English.  
 XX  
 CC The invention relates to a method for screening animals to determine  
 CC those more likely to produce large litters and improved meat quality  
 CC traits. The method involves assaying for the presence of a genotype

CC in the sample of genetic material obtained from animal. The genotype  
CC is characterised by polymorphism(s) in the AMP activated protein  
CC kinase regulatory gamma subunit (PRKAG3) gene. The method is used  
CC for screening animals e.g.; pigs to determine those most likely to  
CC exhibit improved meat quality traits and to produce larger litters.  
CC The present sequence is pig PRKAG3 polymorphic variant (PRKAG3-200).  
XX

SQ Sequence 464 AA;

Query Match 98.68; Score 210; DB 23; Length 464;  
Best Local Similarity 97.64; Pred. No. 1.3e-22;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFRDLA 41  
|||||  
Db 295 HILTHKRLKFLHIFGTLPRPSFLYRTIQDLGIGTFRDLA 335

Search completed: June 6, 2003, 11:01:03  
Job time : 7.60595 secs





```
Query Match      45.5%; Score 97; DB 2; Length 478;
Best Local Similarity 52.8%; Pred. No. 1.1e-05;
Matches 19; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 2 ILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGTF 37
DB 192 ILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGTF 227

RESULT 3
T10971
PV42 protein - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: T10971
R:Abe, H.
submitted to the EMBL Data Library, November 1995
A:Reference number: 217236
A:Accession: T10971
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-379 <ABE>
A:Cross-references: EMBL:U40713; NID:g1113940; PID:g1113941

Query Match      29.6%; Score 63; DB 2; Length 379;
Best Local Similarity 48.6%; Pred. No. 0.54;
Matches 17; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 2 ILTHKLLKFLHIFGSLPRPSFLYRTIQDLGIGT 36
DB 186 MLTQMDMLKFLHGGGAEL--HSLRSKSVQDLGADT 218

RESULT 4
T26720
Hypothetical protein Y45F10A.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T26720
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: 220256
A:Accession: T26720
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1270 <WIL>
A:Cross-references: EMBL:AL021483; PIDN:CAAL6349.1; GSPDB:GN000022; CESP:Y45F10A.6
A:Experimental source: clone Y38H8A
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: 220285
A:Accession: T26911
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1270 <WIL>
A:Cross-references: EMBL:AL021488; PIDN:CAAL16368.1; GSPDB:GN000022; CESP:Y45F10A.6
A:Experimental source: clone Y45F10A
C:Genetics:
A:Gene: CESP:Y45F10A.6
A:Map position: 4
A:Introns: 122/1; 154/1; 183/1; 227/3; 276/2; 324/2; 368/3; 450/3; 701/3; 783/3; 852/3;

Query Match      26.3%; Score 56; DB 2; Length 1270;
Best Local Similarity 43.6%; Pred. No. 20;
Matches 17; Conservative 3; Mismatches 11; Indels 8; Gaps 2;

QY 11 FLHIFGSLPRP---SFLYRT-----IQDLGIGTFRDLA 41
DB 895 FSEVFRLPLPWPNTVFIRVRFLDLISNGLITRDLA 933

RESULT 5
H82145
conserved hypothetical protein VC1874 [Imported] - Vibrio cholerae (strain N16961 se
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82145
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Winn, M.L.; Dodson, F.
charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20408833; PMID:10952301
A:Accession: H82145
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <HEI>
A:Cross-references: GB:AE004263; GB:AE003852; NID:g9656399; PIDN:AAF95022.1; GSPDB:C
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1874
A:Map position: 1

Query Match      26.1%; Score 55.5; DB 2; Length 516;
Best Local Similarity 31.8%; Pred. No. 8.7;
Matches 14; Conservative 9; Mismatches 16; Indels 5; Gaps 1;

QY 2 ILTHKLLKFLHIFGSLPRPSFLYRTIQD-----IGIGTFRD 40
DB 305 LLSDKEMLEFLHSHTSVVAQPPYNSRVFSGINPYALGFAMFERDI 348

RESULT 6
ITHUA2
alpha-2-antiplasmin precursor [validated] - human
N:Alternate names: alpha-2-Pi; alpha-2-plasmin inhibitor precursor
C:Species: Homo sapiens (man)
C>Date: 31-Jul-1989 #sequence_revision 12-Apr-1996 #text_change 08-Dec-2000
C:Accession: A31402; A32163; A41504; A26684; A24708; PC2129; S00068; S32524; S32529
R:Hirotsawa, S.; Nakamura, Y.; Miura, O.; Sumi, Y.; Aoki, N.
Proc. Natl. Acad. Sci. U.S.A. 85, 6836-6840, 1988
A:Title: Organization of the human alpha-2-plasmin inhibitor gene.
A:Reference number: A31402; MUID:88320531; PMID:3166140
A:Accession: A31402
A:Molecule type: DNA
A:Residues: 1-491 <HIRI>
A:Cross-references: GB:M20786; GB:J03830; NID:g177884; PIDN:AAA51554.1; PID:g177886
R:Hirotsawa, S.; Nakamura, Y.; Miura, O.; Sumi, Y.; Aoki, N.
Proc. Natl. Acad. Sci. U.S.A. 85, 1612-1613, 1989
A:Reference number: A32163
A:Accession: A32163
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-491 <HIR2>
R:Tone, M.; Kikuno, R.; Kume-Iwaki, A.; Hashimoto-Gotoh, T.
J. Biochem. 102, 1033-1041, 1987
A:Title: Structure of human alpha-2-plasmin inhibitor deduced from the cDNA sequence
A:Reference number: A41504; MUID:8619254; PMID:2830248
A:Accession: A41504
A:Molecule type: mRNA
A:Residues: 1-32, 'W', 34-491 <TON>
A:Cross-references: GB:D00174; NID:g219409; PIDN:BA00124.1; PID:g219410
R:Holmes, W.E.; Nellies, L.; Liljnen, H.R.; Collen, D.
J. Biol. Chem. 262, 1659-1664, 1987
A:Title: Primary structure of human alpha-2-antiplasmin, a serine protease inhibitor
A:Reference number: A26684; MUID:87109313; PMID:2433286
A:Accession: A26684
A:Molecule type: mRNA
A:Residues: 4-288, 'D', 290-491 <HOL>
A:Cross-references: GB:J02654; NID:g178750; PIDN:AAA35543.1; PID:g178751
A>Note: the authors translated the codon GAT for residue 289 as His
R:Sumi, Y.; Nakamura, Y.; Aoki, N.; Sakai, M.; Muramatsu, M.
J. Biochem. 100, 1399-1402, 1986
A:Title: Structure of the carboxyl-terminal half of human alpha-2-plasmin inhibitor
A:Reference number: A24708; MUID:87137400; PMID:3818581
A:Accession: A24708
```



Query Match Best Local Similarity Matches	<div></div> <div>24.6%; Score 52.5; DB 2; Length 492; 47.2%; Pred. No. 22;</div> <div>Conservative 2; Mismatches 14; Indels 3; Gaps 2;</div>
QY      Dd	4 THKRLLFHLH-FGSLPRPSFLYRTIQDLIGTFR 38 <div>I::I::II                  II 129 TLQRLEKYLHADSGPLPH--LLSRLCQLDGLPGAFR 162</div>
RESULT 11	
S61980	histone acetyltransferase [EC 2.3.1.48], RNA polymerase II-associated [validated] - N/Aternate names: protein LPG22c; protein YPL086c C:/Species: saccharomyces cerevisiae C:/Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 19-Apr-2002 C:/Accession: S61980 R/Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K. submitted to The EMBL Data Library, December 1995 A;/Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A;/Reference number:	S61959
A;/Accession:	S61980
A;/Molecule type:	DNA
A;/Residues:	1-557 <W>
A;/Cross-references:	EMBL:U43281; NID:g1151218; PIDN:AAB68213.1; PID:g1151240; GSPDB:C
C;/Genetics:	Gene: SGD:ELP3; ELP3; MIPS:YPL086c
A;/Cross-references:	SGD:S0006007
A;/Map position:	16L
C;/Function:	
A;/Description:	EC 2.3.1.48 [validated, MUID:99374060]; acetylates core histones involved in chromatin
C;/Superfamily:	hypothetical protein YPL086c
C;/Keywords:	acetyltransferase; coenzyme A; transcription
Query Match Best Local Similarity Matches	<div></div> <div>24.4%; Score 52; DB 2; Length 557; 37.9%; Pred. No. 29;</div> <div>Conservative 7; Mismatches 11; Indels 0; Gaps 0;</div>
QY      Db	8 LLKLHFHGSLRPSPFLYRTIQDLIGIT 36 <div>:::     ::   :     482 IVRELVHYGVSWPLHSRDPRKFQHGGFT 510</div>
RESULT 12	
E87536	ABC transporter, permease protein, probable CC2318 [imported] - Caulobacter crescentus C:/Species: Caulobacter crescentus C:/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001 C;/Accession: E87536 R/Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kohn, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, P., Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001 A;/Title: Complete Genome Sequence of Caulobacter crescentus. A;/Reference number: E87249; MUID:21173698; PMID:11259647
A;/Accession:	E87536
A;/Status:	preliminary
A;/Molecule type:	DNA
A;/Residues:	1-371 <STO>
A;/Cross-references:	GB:AE005673; NID:g13423841; PIDN:AAK24289.1; GSPDB:GN00148
C;/Genetics:	
A;/Gene:	CC2318
Query Match Best Local Similarity Matches	<div></div> <div>24.2%; Score 51.5; DB 2; Length 371; 25.6%; Pred. No. 22;</div> <div>Conservative 11; Mismatches 14; Indels 7; Gaps 1;</div>
QY      Db	2 ILTKRKILKLFHFGSLL-----PPSFLYRTIQDLIGTF 37 <div> :::  : :: :: :: : 77 ILAKQVLRLEDVLGAIVKPQAARTGTVALERIGRVF 119</div>

## RESULT 13

S46088  
hypothetical protein YBR214w - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein YBR1501  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 19-Apr-2002  
C:Accession: S46088; S46090  
R:Rieger, M.  
submitted to the Protein Sequence Database, August 1994

A:Reference number: S45734  
A:Accession: S46088

A:Molecule type: DNA  
A:Residues: 1-527 <RIE>

A:Cross-references: EMBL:Z36083; NID:G536599; PIDN:CAA85178.1; PID:G536600; MIPS:YBR214w  
A:Experimental source: strain S288C  
R:Dubois, E.; El Bakoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.;  
submitted to the Protein Sequence Database, August 1994

A:Reference number: S45782  
A:Accession: S46090

A:Molecule type: DNA  
A:Residues: 114-527 <DUB>

A:Cross-references: EMBL:Z36083; MIPS:YBR214w  
A:Experimental source: strain S288C

C:Genetics:

A:Gene: SCD:SDS24

A:Cross-references: SCD:S0000418

A:Map position: 2R

Query Match 24.2%; Score 51.5; DB 2; Length 527;

Best Local Similarity 42.1%; Pred. No. 33;

Matches 16; Conservative 6; Mismatches 9; Indels 7; Gaps 2;

QY 2 ILTHKRLKEL----HIFGSLPRPSFLYRTIIDLIGT 35

||: |||:| | | | | | | | | | | | | |

Db 240 ILSQRRLIKYLDNARSFTSLEP---LLNSQLDLHIG 274

## RESULT 14

T18227

hypothetical protein - yeast (*Candida albicans*)

C:Species: *Candida albicans*

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T18227

R:Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z18831

A:Accession: T18227

A:Status: preliminary; translated from GB/EMBL/NDRI

A:Molecule type: DNA

A:Residues: 1-629 <BAR>

A:Cross-references: EMBL:AL033501; NID:el341022; PID:el341023; PIDN:CAA21985.1

C:Genetics:

A:Note: Ca41C10.01c

Query Match

Best Local Similarity 24.2%; Score 51.5; DB 2; Length 629;

Matches 14; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 2 ILTHKRLKELHIFGSLPRPSF-LYRTIIDLIGT 36

||: |||:| | | | | | | | | | | | | |

Db 300 ILSQRRLIKYMNARRRPFSLDFYLTSTLQDLKIGS 335

## RESULT 15

C64677

conserved hypothetical protein HPI259 - *Helicobacter pylori* (strain 26695)

C:Species: *Helicobacter pylori*

C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 11-Jan-2000

C:Accession: C64677

R:Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frase  
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: C64677

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-205 <TOM>

A:Cross-references: GB:AE000631; GB:AE000511; NID:q2314421; PIDN:AAD08305.1; PID:g2  
C:Superfamily: conserved hypothetical protein b1120

Query Match 23.9%; Score 51; DB 2; Length 205;

Best Local Similarity 37.8%; Pred. No. 13;

Matches 14; Conservative 6; Mismatches 9; Indels 8; Gaps 2;

QY 10 KFLHIFGSL-----PRPSFLYRTIIDLIGTFRDLA 41

: ||: | || | | | | | | | | | | | | |

Db 75 RILHLHSELLSVRSEKDPNLVYRNEKDLNIG---DLA 108

Search completed: June 6, 2003, 11:03:38

Job time : 4.65674 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:57:04 ; Search time 1.50788 Seconds  
(without alignments)  
1127.761 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_320\_360

Perfect score: 213

Sequence: 1 HLTHKRLKFLHIFGSLP.....PSFLYRTIQDLGIGTRDLA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	213	100.0	464	1 AAKI_HUMAN	Q9UG19 homo sapien
2	210	98.6	464	1 AAKI_PIG	Q9MYP4 sus scrofa
3	124	58.2	569	1 AAKE_HUMAN	Q9UGJ0 homo sapien
4	112	52.6	330	1 AAKG_BOVIN	P58108 bos taurus
5	112	52.6	330	1 AAKG_MOUSE	O54950 mus musculus
6	112	52.6	330	1 AAKG_RAT	P80385 rattus norv
7	112	52.6	331	1 AAKG_HUMAN	P54619 homo sapien
8	64.5	30.3	133	1 AAKG_PIG	Q09138 sus scrofa
9	54.5	25.6	491	1 A2AP_HUMAN	P08697 homo sapien
10	54	25.4	269	1 YD55_SCHPO	Q10308 schizosacch
11	53	24.9	305	1 HEM6_PSEAE	P43898 pseudomonas
12	53	24.9	1024	1 S26L_HUMAN	Q9BYH1 homo sapien
13	52.5	24.6	492	1 A2AP_BOVIN	P28800 bos taurus
14	51.5	24.2	527	1 YB64_YEAST	P38314 saccharomyc
15	51	23.9	309	1 HEM6_YERPE	Q62CF9 yersinia pe
16	50	23.5	1167	1 ITAE_MOUSE	Q60677 mus musculus
17	49.5	23.2	606	1 ABD4_MOUSE	O89016 mus musculus
18	49.5	23.2	2136	1 YCF2_MARPO	P09975 marchantia
19	49	23.0	304	1 LST_HAEIN	O48211 haemophilus
20	49	23.0	712	1 TRAL_HUMAN	P51617 homo sapien
21	48.5	22.8	1786	1 UVRA_CHLTR	O84337 chlamydia t
22	48.5	22.8	1840	1 SUIS_RAT	P23739 rattus norv
23	48	22.5	186	1 Y786_METJA	Q58196 methanococc
24	48	22.5	210	1 YD07_HAEIN	O57320 haemophilus
25	48	22.5	320	1 MTDL_YEAST	Q02046 saccharomyc
26	48	22.5	420	1 Y4ME_RHISN	P55564 rhizobium s
27	47.5	22.3	494	1 CAT4_PEA	P25890 pisum sativ
28	47	22.1	271	1 LCRF_YERPE	P28808 yersinia pe
29	47	22.1	271	1 VIRE_YEREN	P13225 yersinia en
30	47	22.1	454	1 ATTY_RAT	P04694 rattus norv
31	47	22.1	798	1 CDB2_HUMAN	O9Y5E7 homo sapien
32	46.5	21.8	161	1 CRAA_TRIIN	P02500 trichechus
33	46.5	21.8	353	1 BRBI_HUMAN	P46663 homo sapien

34	46.5	21.8	606	1 ABD4_HUMAN	O14678 homo sapien
35	46	21.6	184	1 RR4_PATFR	O20273 paterosonia
36	46	21.6	196	1 RR4_PATSQ	O36052 paterosonia
37	46	21.6	307	1 YC23_SYNY3	P73467 synecocyst
38	46	21.6	346	1 TH13_SCHPO	P36597 schizosacch
39	46	21.6	471	1 SECF_MYCLE	P38386 mycobacteri
40	46	21.6	471	1 YODO_BACSU	O34676 bacillus su
41	46	21.6	692	1 EFG_MYCPU	O98QD8 mycoplasma
42	46	21.6	847	1 SYA_HELPJ	O92JY5 helicobacte
43	46	21.6	847	1 SYA_HELPY	P56452 helicobacte
44	46	21.6	923	1 PH87_YEAST	P25360 saccharomyc
45	46	21.6	1036	1 Y414_MYCGE	P47653 mycoplasma

#### ALIGNMENTS

RESULT 1  
AAKI\_HUMAN STANDARD: PRT; 464 AA.  
ID Q9UG19; Q9NRLI;  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)  
DE (AMPK gamma3).  
GN PRKAG3 OR AMPK3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20164049; PubMed=10698692;  
RA Cheung P.C., Salt I.P., Davies S.P., Hardie D.G., Carling D.;  
RT "Characterization of AMP-activated protein kinase gamma-subunit  
isoforms and their role in AMP binding.";  
RL Biochem. J. 346:659-669(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=20280150; PubMed=10818001;  
RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,  
Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,  
Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,  
Andersson L.;  
RT "A mutation in PRKAG3 associated with excess glycogen content in pig  
skeletal muscle.";  
RL Science 288:1248-1251(2000).  
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE  
IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.  
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
CC -!- TISSUE SPECIFICITY: SKELETAL MUSCLE, WITH WEAK EXPRESSION IN HEART  
AND PANCREAS.  
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
SUBUNIT FAMILY.  
CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; AJ249977; CAB65117.1; ALT\_INIT.  
DR EMBL; AF214519; AAF73987.1; -;

```

DR Genew: HGNC:9187; PRKAG3.
DR MIX: 604976; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
DR Fatty acid biosynthesis; Repeat; CBS domain.
DR FT DOMAIN 147 201 CBS 1.
DR FT DOMAIN 228 282 CBS 2.
DR FT DOMAIN 303 356 CBS 3.
DR FT DOMAIN 375 428 CBS 4.
DR FT CONFLICT 58 58 T -> A (IN REF. 1).
DR FT CONFLICT 163 164 MQ -> IE (IN REF. 1).
DR FT CONFLICT 398 398 Q -> K (IN REF. 1).
DR FT CONFLICT 461 464 ALGA -> PSGPEKI (IN REF. 1).
DR FT CONFLICT 464 464 ALGA -> PSGPEKI (IN REF. 1).
DR SQ SEQUENCE 464 AA; 51514 MW; 53985C2C77003A63 CRC64;

Query Match 100.0%; Score 213; DB 1; Length 464;
Best Local Similarity 100.0%; Pred. No. 2.1e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB 295 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 2
AAKI_PIG
ID AAKI_PIG STANDARD; PRT; 464 AA.
AC Q9WY24;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)
DE (AMPK gamma3).
GN PRKAG3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT RN(-) GLN-200.
RC TISSUE-Skeletal muscle.
RX MEDLINE=20280150; PubMed=10818001;
RA Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,
RA Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H.,
RA Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,
RA Andersson L.;
RA "A mutation in PRKAG3 associated with excess glycogen content in pig
RT skeletal muscle.";
RL Science 288:1248-1251(2000).
CC -1- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. IT MAY PLAY A ROLE
CC IN THE REGULATION OF ENERGY METABOLISM IN SKELETAL MUSCLE.
CC -1- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -1- TISSUE SPECIFICITY: MUSCLE.
CC -1- DISEASE: DEFECTS IN PRKAG3 (RN-) ARE ASSOCIATED WITH EXCESS
CC GLYCOGEN CONTENT (ABOUT 70%) IN SKELETAL MUSCLE. THIS MUTATION
CC ORIGINATED IN THE HAMPSHIRE BREED PIGS AND HAS BENEFICIAL EFFECTS
CC ON MEAT CONTENT BUT DETRIMENTAL EFFECTS ON PROCESSING YIELD. THUS,
CC THIS MUTATION IS OF CONSIDERABLE ECONOMIC SIGNIFICANCE IN THE PIG
CC BREEDING INDUSTRY.
CC -1- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF214521; AAF73988.1; -.
DR EMBL: AF214520; AAF73988.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
DR Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation.
DR FT DOMAIN 172 226 CBS 1.
DR FT DOMAIN 253 307 CBS 2.
DR FT DOMAIN 328 381 CBS 3.
DR FT DOMAIN 400 453 CBS 4.
DR FT VARIANT 200 200 R -> Q (IN RN-).
DR SQ SEQUENCE 464 AA; 51308 MW; 17638CB12A2BA9DF CRC64;

Query Match 98.6%; Score 210; DB 1; Length 464;
Best Local Similarity 97.6%; Pred. No. 5.4e-21;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB 295 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 335

RESULT 3
AAKH_HUMAN
ID AAKH_HUMAN STANDARD; PRT; 569 AA.
AC Q9UGJ0; Q9UGJ0; Q9NU29; Q9ULX8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-2 subunit (AMPK gamma-2 chain)
DE (AMPK gamma2) (H91620P).
GN PRKAG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=20164049; PubMed=10698692;
RA Cheung P.C.F., Salt I.P., Davies S.P., Hardie D.G., Carling D.;
RT "Characterization of AMP-activated protein kinase gamma-subunit
RT isoforms and their role in AMP binding.";
RL Biochem. J. 346:659-669(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM B).
RX MEDLINE=20564210; PubMed=11112354;
RA Lang T., Yu L., Qiang T., Jiang J., Chen Z., Xin Y., Liu G., Zhao S.;
RT "Molecular cloning, genomic organization, and mapping of PRKAG2, a
RT heart abundant gamma-2 subunit of 5'-AMP-activated protein kinase, to
RT human chromosome 7q36.";
RL Genomics 70:258-263(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 218-569 FROM N.A.
RA Hattori A., Seki N., Hayashi A., Kozuma S., Muramatsu M., Saito T.;
RT "Human homolog of AMPK gamma-1 chain.";

```



```

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Benkel B., Kollers S., Fries R., Sazanov A., Yoshida E., Davoren J.,
RA Hickey D.;
RT "Characterization of the bovine AMPK gamma-1 gene.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF329081; AAK19307.1; -
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 48 96 CBS 1.
FT DOMAIN 123 177 CBS 2.
FT DOMAIN 197 250 CBS 3.
FT DOMAIN 271 323 CBS 4.
SQ SEQUENCE 330 AA; 37481 MW; F130AC8EE2BEF89 CRC64;

Query Match 52.6%; Score 112; DB 1; Length 330;
Best Local Similarity 46.3%; Pred. No. 7.3e-08;
Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB 165 YILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 205

RESULT 5
AAKG_MOUSE STANDARD; PKT; 330 AA.
AC O34950;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPK)
GN PRKAG1 OR PRKAAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21203559; PubMed=11306812;
RA Shamsadin R., Jantsan K., Adham I., Engel W.;
RT "Cloning, organisation, chromosomal localization and expression
RT analysis of the mouse Prkag1 gene.";
RL Cytogenet. Cell genet. 92:134-138(2001).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.

```

```

CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A
CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF036535; AAB95475.1; -
DR MGD; MGI:108411; Prkag1
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 47 95 CBS 1.
FT DOMAIN 122 176 CBS 2.
FT DOMAIN 196 249 CBS 3.
FT DOMAIN 270 322 CBS 4.
SQ SEQUENCE 330 AA; 37554 MW; D504707B83512DDC CRC64;

Query Match 52.6%; Score 112; DB 1; Length 330;
Best Local Similarity 48.8%; Pred. No. 7.3e-08;
Matches 20; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 HILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
DB 164 YILTHKRLKFLHFGSLPRPSFLYRTIQDLGIGTFRDLA 204

RESULT 6
AAKG_RAT STANDARD; PKT; 330 AA.
AC P80385;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPK)
GN PRKAG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=96215327; PubMed=8626596;
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,
RA Beri R.K., Carling D.;
RT "Characterization of AMP-activated protein kinase beta and gamma
RT subunits. Assembly of the heterotrimeric complex in vitro.";
RL J. Biol. Chem. 271:10282-10290(1996).
RN [2]
RP SEQUENCE OF 8-330 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT "Non-catalytic beta- and gamma-subunit isoforms of the
RT 5'-AMP-activated protein kinase.";
RL J. Biol. Chem. 271:8675-8681(1996).
RN [3]
RP SEQUENCE OF 48-330 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;

```



RT Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are  
 RL homologs of proteins that interact with yeast Snf1 protein kinase.";  
 CC J. Biol. Chem. 269:29343-29346(1994).  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HORMONE-SENSITIVE LIPASE AND  
 CC HYDROXYMETHYLGLUTARYL-COA REDUCTASE. APPEARS TO ACT AS A METABOLIC  
 CC STRESS-SENSING PROTEIN KINASE SWITCHING OFF BIOSYNTHETIC PATHWAYS  
 CC WHEN CELLULAR ATP LEVELS ARE DEPLETED AND WHEN 5'-AMP RISES IN  
 CC RESPONSE TO FUEL LIMITATION AND/OR HYPOXIA. THIS IS A REGULATORY  
 CC SUBUNIT.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNIT.  
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART AND BRAIN, ALSO  
 CC FOUND IN KIDNEY, WHITE ADIPOSE TISSUE, LUNG AND SPLEEN.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; X95578; CAA64831.1; -.  
 DR EMBL; U24113; AAC52580.1; -.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam; PF00571; CBS; 4.  
 DR SMART; SM00116; CBS; 4.  
 DR Fatty acid biosynthesis; Repeat; CBS domain.  
 KW DOMAIN 127 95 CBS 1.  
 FT DOMAIN 122 176 CBS 2.  
 FT DOMAIN 196 249 CBS 3.  
 FT DOMAIN 270 322 CBS 4.  
 FT CONFLICT 114 114 E->Q (IN REF. 3).  
 FT CONFLICT 201 201 A->P (IN REF. 3).  
 SQ SEQUENCE 330 AA; 37386 MW; 36031E526C1F1E97 CRC64;  
 Query Match 52.6%; Score 112; DB 1; Length 330;  
 Best Local Similarity 46.3%; Pred. No. 7.3e-08;  
 Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 HILTHRLKLFHIFGSLPRPSFLYRTQDLCIGTFRDLA 41  
 DB 164 YILTHRLKLFKLFITEFPKPFMSKLSLEELQIGTYANIA 204  
 RESULT 7  
 AAKG\_HUMAN STANDARD; PRT; 331 AA.  
 AC P54619;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)  
 DE (AMPK).  
 GN PRKAG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=96224074; PubMed=8621499;  
 RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,  
 RA Dyck J.R.B., Kemp B.E., Witters L.A.;  
 RT "Non-catalytic beta- and gamma-subunit isoforms of the 5'-AMP-activated  
 RT protein kinase.";

RL J. Biol. Chem. 271:8675-8681(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
 CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNIT.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; U24112; AAC50495.1; -.  
 DR EMBL; BC000358; AAH00358.1; -.  
 DR Genew; HGNC:9385; PRKAG1.  
 DR MIM; 602742; -.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam; PF00571; CBS; 4.  
 DR SMART; SM00116; CBS; 4.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT DOMAIN 123 177 CBS 1.  
 FT DOMAIN 197 250 CBS 2.  
 FT DOMAIN 271 323 CBS 3.  
 FT DOMAIN 271 323 CBS 4.  
 SQ SEQUENCE 331 AA; 37579 MW; 0F22B9CALDBD87AE CRC64;  
 Query Match 52.6%; Score 112; DB 1; Length 331;  
 Best Local Similarity 46.3%; Pred. No. 7.3e-08;  
 Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 HILTHRLKLFHIFGSLPRPSFLYRTQDLCIGTFRDLA 41  
 DB 165 YILTHRLKLFKLFITEFPKPFMSKLSLEELQIGTYANIA 205  
 RESULT 8  
 AAKG\_PIG STANDARD; PRT; 133 AA.  
 AC Q09138;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)  
 DE (AMPK) (38 kDa subunit) (Fragments).  
 GN PRKAG1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=95050763; PubMed=7961907;  
 RA Stapleton D., Gao G., Mitchell B.J., Widmer J., Mitchell K.,  
 RA Teh T., House C.M., Witters L.A., Kemp B.E.;  
 RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are  
 RT homologs of proteins that interact with yeast Snf1 protein kinase.";  
 RL J. Biol. Chem. 269:29343-29346(1994).  
 CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
 CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO

CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
 CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
 CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.  
 CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA CATALYTIC SUBUNIT, A BETA AND A  
 CC GAMMA NON-CATALYTIC REGULATORY SUBUNITS.  
 CC -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA  
 CC SUBUNIT FAMILY.  
 CC -!- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 DR InterPro: IPR000644; CBS\_domain.  
 DR Pfam: PF00571; CBS; 1.  
 DR SMART: SM00116; CBS; 1.  
 KW Fatty acid biosynthesis; Repeat; CBS domain.  
 FT NON\_TER 1  
 FT DOMAIN <1 >11 CBS 1.  
 FT DOMAIN <26 >42 CBS 2.  
 FT DOMAIN 62 91 CBS 3.  
 FT DOMAIN <95 127 CBS 4.  
 FT NON\_CONS 11 12  
 FT NON\_CONS 24 25  
 FT NON\_CONS 42 43  
 FT NON\_CONS 73 74  
 FT NON\_CONS 80 81  
 FT NON\_CONS 94 95  
 FT NON\_CONS 103 104  
 FT NON\_TER 133 133  
 SQ SEQUENCE 133 AA; 14763 MW; 84C0C3D41E845CEF CRC64;  
 Query Match 30.3%; Score 64.5; DB 1; Length 133;  
 Best Local Similarity 31.7%; Pred. No. 0.076;  
 Matches 13; Conservative 13; Mismatches 8; Indels 7; Gaps 1;  
 QY 1 HILTHRLKLFHFGSLPRPFLYTKTQDLGIGTFRDLA 41  
 DB 37 YILTXK-----LFITEPKPEFMSKLEELQIGTYANIA 70  
 RESULT 9  
 A2AP HUMAN STANDARD; PRT; 491 AA.  
 AC P08697;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alpha-2-antiplasmin precursor (Alpha-2-plasmin inhibitor) (Alpha-2-PI)  
 GN (Alpha-2-AP).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88139254; PubMed=2830248;  
 RA Tone M., Kikuno R., Kume-Iwaki A., Hashimoto-Gotoh T.;  
 RT "Structure of human alpha 2-plasmin inhibitor deduced from the cDNA  
 RT sequence.";  
 RL J. Biochem. 102:1033-1041(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88320531; PubMed=3166140;  
 RA Hirotsawa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;  
 RT "Organization of the human alpha 2-plasmin inhibitor gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6836-6840(1988).  
 RN [3]  
 RP ERRATUM.  
 RA Hirotsawa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:1612-1613(1989).  
 RN [4]  
 RP SEQUENCE OF 4-491 FROM N.A.  
 RX MEDLINE=87109313; PubMed=2433286;  
 RA Holmes W.E., Nelles L., Lijnen H.R., Collen D.;  
 RT "Primary structure of human alpha 2-antiplasmin, a serine protease  
 RT inhibitor (serpin).";

RL J. Biol. Chem. 262:1659-1664(1987).  
 RN [5]  
 RP SEQUENCE OF 218-491 FROM N.A.  
 RX MEDLINE=87137400; PubMed=3818581;  
 RA Sumi Y., Nakamura Y., Aoki N., Sakai M., Murematsu M.;  
 RT "Structure of the carboxyl-terminal half of human alpha 2-plasmin  
 RT inhibitor deduced from that of cDNA.";  
 RL J. Biochem. 100:1399-1402(1986).  
 RN [6]  
 RP SEQUENCE OF 40-491.  
 RX MEDLINE=87275946; PubMed=2440681;  
 RA Lijnen H.R., Holmes W.E., van Hoef B., Wiman B., Rodriguez H.,  
 RA Collen D.;  
 RT "Amino-acid sequence of human alpha 2-antiplasmin.";  
 RL Eur. J. Biochem. 166:565-574(1987).  
 RN [7]  
 RP SEQUENCE OF 40-43.  
 RX MEDLINE=78023887; PubMed=21075;  
 RA Wiman B., Collen D.;  
 RT "Purification and characterization of human antiplasmin, the  
 RT fast-acting plasmin inhibitor in plasma.";  
 RL Eur. J. Biochem. 78:19-26(1977).  
 RN [8]  
 RP SEQUENCE OF 28-52.  
 KC TISSUE=Plasma;  
 RX MEDLINE=93050153; PubMed=1385210;  
 RA Christensen S., Sottrup-Jensen L.;  
 RT "Bovine alpha 2-antiplasmin. N-terminal and reactive site sequence.";  
 RL FEBS Lett. 312:100-104(1992).  
 RN [9]  
 RP ACTIVE SITES  
 RX MEDLINE=88290696; PubMed=2456616;  
 RA Potempa J., Shieh B.-H., Travis J.;  
 RT "Alpha-2-antiplasmin: a serpin with two separate but overlapping  
 RT reactive sites.";  
 RL Science 241:699-700(1988).  
 RN [10]  
 RP SEQUENCE OF 481-491, AND SULFATON.  
 RX MEDLINE=87137577; PubMed=2434496;  
 RA Hortin G., Fok K.F., Toren P.C., Strauss A.W.;  
 RT "Sulfation of a tyrosine residue in the plasmin-binding domain of  
 RT alpha 2-antiplasmin.";  
 RL J. Biol. Chem. 262:3082-3085(1987).  
 RN [11]  
 RP VARIANT OKINAWA.  
 RX MEDLINE=90036902; PubMed=2572590;  
 RA Miura O., Sugahara Y., Aoki N.;  
 RT "Hereditary alpha 2-plasmin inhibitor deficiency caused by a  
 RT transport-deficient mutation (alpha 2-PI-Okina) Deletion of Glu137  
 RT by a trinucleotide deletion blocks intracellular transport.";  
 RL J. Biol. Chem. 264:18213-18219(1989).  
 RN [12]  
 RP VARIANT ALPHA-2-PLASMIN INHIBITOR DEFICIENCY M-411, AND VARIANTS V-27;  
 RP W-33 AND K-434.  
 RX MEDLINE=20051147; PubMed=10583218;  
 RA Lind B., Thorsen S.;  
 RT "A novel missense mutation in the human plasmin inhibitor  
 RT (alpha2-antiplasmin) gene associated with a bleeding tendency.";  
 RL Br. J. Haematol. 107:317-322(1999).  
 CC -!- FUNCTION: THE MAJOR TARGETS OF THIS INHIBITOR ARE PLASMIN AND  
 CC TRYPSIN, BUT IT ALSO INACTIVATES CHYMOTRYPSIN.  
 CC -!- DISEASE: Defects in SERPINF2 are the cause of alpha-2-plasmin  
 CC inhibitor deficiency, a disease resulting in severe hemorrhagic  
 CC diathesis.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

-----
CC EMBL; D00116; BRA00070.1; -.
DR EMBL; D00174; BRA00124.1; -.
DR EMBL; M20786; AA51554.1; -.
DR EMBL; M20782; AA51554.1; JOINED.
DR EMBL; M20783; AA51554.1; JOINED.
DR EMBL; M20784; AA51554.1; JOINED.
DR EMBL; M20785; AA51554.1; JOINED.
DR EMBL; J02654; AA33543.1; -.
DR EMBL; A24708; A24708.
DR PIR; A26684; A26684.
DR PIR; A31402; A31402.
DR PIR; A32163; A32163.
DR PIR; A41504; A41504.
DR PIR; S00068; S00068.
DR HSP; O35684; IJJO.
DR SWISS-2DPAGE; P08697; HUMAN.
DR Genew; HGNC:9075; SERPIN2.
DR MIM; 262850; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma; Signal;
KW Sulfation; Polymorphism; Disease mutation.
FT SIGNAL 1 27
FT PROPEP 28 39
FT CHAIN 40 491
FT ACT_SITE 403 404
FT ACT_SITE 404 405
FT MOD_RES 484 484
FT CARBOHYD 126 126
FT CARBOHYD 295 295
FT CARBOHYD 309 309
FT CARBOHYD 316 316
FT VARIANT 27 27
FT VARIANT 33 33
FT VARIANT 176 176
FT VARIANT 411 411
FT VARIANT 434 434
FT VARIANT 434 434
FT CONFLICT 49 49
FT CONFLICT 105 105
FT CONFLICT 289 289
FT CONFLICT 408 408
FT CONFLICT 455 455
SQ SEQUENCE 491 AA; 54565 MW; 385A1C90E91A63BC CRC64;

Query Match 25.6%; Score 54.5; DB 1; Length 491;
Best Local Similarity 48.6%; Pred. No. 7;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 4 THKRLKLFHFGSLPRPSFLYRTIQDLGIGTFR 38
Db 128 TLQRLQVLAH-GSGPCLPHLLSLRCQDLGPGAFR 161

RESULT 10
YD55_SCHPO
ID YD55_SCHPO STANDARD; PRT; 269 AA.
AC Q10308;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C6C3.05 in chromosome I.

```

```

GN SPAC6C3.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam K., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; 269731; CAB40280.1; -.
KW Hypothetical protein.
SQ SEQUENCE 269 AA; 31234 MW; 0C3DF87A138CA5BE CRC64;

Query Match 25.4%; Score 54; DB 1; Length 269;
Best Local Similarity 35.9%; Pred. No. 4.3;
Matches 14; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

Qy 2 ILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTFRDL 40
Db 216 IKTHYKTLRIHTLTLFPVPSVLSSNLKQMGWFEFL 254

RESULT 11
HEM6_PSEAE
ID HEM6_PSEAE STANDARD; PRT; 305 AA.
AC P43898;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMF OR PA0024.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RA Hungerer C., Troup B., Jahn D.;
RT "Cloning and regulation of the Pseudomonas aeruginosa hemf gene
RL encoding oxygen-dependent coproporphyrinogen III oxidase.";
RN Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RL opportunistic pathogen.";
RN Nature 406:959-964(2000).
CC -!- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC protoporphyrinogen-IX + 2 CO(2).
CC -!- COFACTOR: IRON (By SIMILARITY).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85015; CAA59376.1; .
DR EMBL; AE004442; AAG03414.1; .
DR EMBL; IPR001260; Coprogen.Oxidas.
DR Pfam: PF01218; Coprogen.Oxidas; 1.
DR PRINTS: PR00073; COPROGNXOXDASE.
DR PROSITE: PS01021; COPROGEN_OXIDASE; 1.
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.
SQ SEQUENCE 305 AA; 34806 MW; 240BD06FEA37EAE3 CRC64;

Query Match 24.9%; Score 53; DB 1; Length 305;
Best Local Similarity 36.8%; Pred. No. 6.7;
Matches 14; Conservative 3; Mismatches 15; Indels 6; Gaps 1;

QY 3 LTHKRLKLFHIFGSLLP-----RPSFLYRTIQLGI 34
| | | | | | | | | | | | | | | | | |
Db 55 LIEKGGVNSHVGDSLPPSASARPELAGRGFQALGV 92
| | | | | | | | | | | | | | | | | |

RESULT 12
S26L.HUMAN STANDARD; PRT; 1024 AA.
AC Q9BYH1; Q9Y2EL; Q9NU13; Q9NU14; Q9NU15; Q9Y917; Q9Y336;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Seizure 6-like protein precursor.
GN SE26L OR KIAA0927.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Brain;
RX PubMed=11175339;
RA Nishioka M., Kohno T., Takahashi M., Niki T., Yamada T., Sone S.,
RA Yokota J.;
RT "Identification of a 428-kb homozygously deleted region disrupting the

```

```

RT SE26L gene at 22q12.1 in a lung cancer cell line.";
RL Oncogene 19:6251-6260(2000).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Brain;
RX MEDLINE=99245063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro.";
RN DNA Res. 6:63-70(1999).
[3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fetal brain;
RA Collins J.E., Huckle E.J.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagdoley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Evans P.D., Dockree C., Dodsorth S.J., Durbin R.M., Ellington A.G.,
RA Dhani K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudon J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Korf I., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Seroussi E., Franssion I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
[5]
RP GENE EXPRESSION REGULATION.
RN PubMed=11992124;
RX Suzuki H., Gabrielson E., Chen W., Anbazhagan R., Van Engeland M.,
RA Weitenberg M.P., Herman J.G., Baylin S.B.;
RT "A genomic screen for genes upregulated by demethylation and histone
RT deacetylase inhibition in human colorectal cancer.";
RL Nat. Genet. 31:141-149(2002).
CC -!- FUNCTION: Candidate tumor suppressor gene.
CC -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1, 2, 3 and 4 (shown here);

```



```
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 28 28 T -> Q (IN REF. 2).
FT CONFLICT 40 40 Q -> P (IN REF. 2).
FT CONFLICT 43 43 Q -> E (IN REF. 2).
SQ SEQUENCE 492 AA; 54710 MW; 0755D6FC89B2DF5D CRC64;

Query Match 24.6%; Score 52.5; DB 1; Length 492;
Best Local Similarity 47.2%; Pred. No. 13;
Matches 17; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

QY 4 THRLKFLKFLHI-FQSLPRPSFLYRTIQDLGIGTFR 38
DB 129 TLORKEVLHADSGLPH--LLSRCLQDLGPGAFR 162

RESULT 14
YB64_YEAST
ID YB64_YEAST STANDARD; PRT; 527 AA.
AC P38314;1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 36, Last annotation update)
DE Hypothetical 57.2 kDa protein in MET8-HPC2 intergenic region.
GN YB214W OR YBR1501.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN Rieger M.;
RP SEQUENCE FROM N.A.
RC STRAIN-S288c;
RA Rieger M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DDJB databases.
[2]
RP SEQUENCE OF 114-527 FROM N.A.
RC STRAIN-S288c;
RA Dubois E., el Bakkoury M., Glansdorff N., Messenguy F., Pierard A.,
RA Scherens B., Vierendeels F.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DDJB databases.
CC -!- SIMILARITY: TO YEAST YGL056C AND S.POMBE SDS23.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z36083; CAAB5178.1; -.
DR PIR; S46088; S46088.
DR SGD; S00004118; SDS24.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 2.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 196 251 CBS 1.
FT DOMAIN 283 335 CBS 2.
SQ SEQUENCE 527 AA; 57187 MW; DC2741550A69C154 CRC64;

Query Match 24.2%; Score 51.5; DB 1; Length 527;
Best Local Similarity 42.1%; Pred. No. 19;
Matches 16; Conservative 6; Mismatches 9; Indels 7; Gaps 2;

QY 2 ILTHKLLKFL----HIFGSLPRPSFLYRTIQDLGIG 35
DB 240 ILSQRLLIKYLDNRARSFTSLEP---LLNSQLDLHG 274

RESULT 15
```

```
HEM6_YERPE
ID HEM6_YERPE STANDARD; PRT; 309 AA.
AC Q82CF9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic (ec 1.3.3.3)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMF OR YPO3032.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feilwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -!- CARBATIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC protoporphyrinogen-IX + 2 CO(2).
CC -!- COFACTOR: Iron (By similarity).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ414155; CAC92274.1; -.
DR InterPro; IPR001260; Coprogen_oxidase.
DR Pfam; PF01218; Coprogen_oxidase; 1.
DR PRINTS; PR00073; COPRGNOXDASE.
DR PROSITE; PS01021; COPROGEN_OXIDASE; 1.
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.
SQ SEQUENCE 309 AA; 35004 MW; A3DE259C4C3714D9 CRC64;

Query Match 23.9%; Score 51; DB 1; Length 309;
Best Local Similarity 34.4%; Pred. No. 13;
Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

QY 9 LKFLHIFGSLP-----RPSFLYRTIQDLGI 34
DB 62 VNFSHVSGAMLPASATARPDELGRSFQALGV 93

Search completed: June 6, 2003, 11:01:31
Job time : 2.50788 secs
```

RESULT 2		
Q9BHL6	PRELIMINARY;	PRT; 372 AA.
ID	Q9BHL6	
AC	Q9BHL6;	
DT	01-JUN-2001	(TREMBLrel. 17, Created)
DT	01-JUN-2001	17, Last sequence up
DT	01-JUN-2001	(TREMBLrel. 21, Last annotation
DT	01-JUN-2002	21, Last annotation

RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T., RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., RA Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A., Yeh J., Yeh R.-F., Zavari J.S., Zhan M., Zhang G., Zhao O., Zheng L., RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; RT "The genome sequence of Drosophila melanogaster."; RL Science 287:2185-2195(2000).
DR	EMBL; AE003733; AAF5864.1; -.
DR	FlyBase; FBgn0025803; SNF4A-gamma.
DR	InterPro; IPR000644; CBS_domain.
DR	Pfam; PF00571; CBS; 4.
DR	SMART; SW00116; CBS; 4.
SQ	SEQUENCE 634 AA; 70174 MW; 3FFD0D53E54BBE7C CRC64;
Query Match 49.3%; Score 105; DB 5; Length 634;	
Best Local Similarity 40.0%; Pred.No. 2.9e-06;	
Matches 16; Conservative 18; Mismatches 6; Indels 0; Gaps	
QY	1 HILTHKRLLKHLFGSLPRSPFLYRTIDLGIGTFRD 40 :     :-  :::    ::: ::   :    :
DB	293 YILTHKRILRFELYINELPKPAYWQSRLKLTGTYNNI 332 :     :-  :::    ::: ::   :    :
RESULT 4	
O96613	ID O96613 PRELIMINARY; PRT; 647 AA.
AC	O96613;
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	SNF4/AMP-activated protein kinase gamma subunit.
OS	SNF4A-GAMMA OR CGI7399.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
ON	[1]
RP	SEQUENCE FROM N.A.
RQ	Yoshida E.N., Benkel B.F., Fong Y., Hickey D.A.;
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF094764; AAC95306.1; -.
DR	EMBL; AF094763; AAC95305.1; -.
DR	FlyBase; FBgn0025803; SNF4A-gamma.
DR	InterPro; IPR000644; CBS_domain.
DR	Pfam; PF00571; CBS; 4.
DR	SMART; SW00116; CBS; 4.
KW	Kinase.
SQ	SEQUENCE 647 AA; 71592 MW; B792BE1089730B52 CRC64;
Query Match 49.3%; Score 105; DB 5; Length 647;	
Best Local Similarity 40.0%; Pred.No. 3e-06;	
Matches 16; Conservative 18; Mismatches 6; Indels 0; Gaps	
QY	1 HILTHKRLLKHLFGSLPRSPFLYRTIDLGIGTFRD 40 :     :-  :::    ::: ::   :    :
DB	306 YILTHKRILRFELYINELPKPAYWQSRLKLTGTYNNI 345 :     :-  :::    ::: ::   :    :
RESULT 5	
Q8SXT8	ID O8SXT8 PRELIMINARY; PRT; 906 AA.
AC	O8SXT8;
DT	01-JUN-2002 (TrEMBLrel. 21, Created)
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)



DE	Hypothetical 47.5 kDa protein.	
GN	T20F7.6.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
QC	Rhabdiida; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
NP	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RX	MEDLINE=99069613; PubMed=9851916;	
RA	None;	
RT	"Genome sequence of the nematode C. elegans: a platform for	
KT	investigating biology. The C. elegans Sequencing Consortium.";	
RL	Science 282:2012-2018(1998).	
RP	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RA	Miller N., Gattung S.;	
RT	"The sequence of C. elegans cosmid T20F7.";	
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.	
RP	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RA	Waterston R.;	
RT	"Direct Submission.";	
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; U97550; AAK18981.2; -.	
DR	InterPro: IPR000644; CBS_domain.	
DR	Pfam: PF00571; CBS; 4.	
DR	SMART; SM00116; CBS; 4.	
KW	Hypothetical protein.	
SK	SEQUENCE 423 AA; 47451 MW; 6997065D515E7B21 CRC64;	
QY	Query Match 45.5%; Score 97; DB 5; Length 423;	
	Best Local Similarity 52.8%; Pred. No. 2.6e-05;	
	Matches 19; Conservative 9; Mismatches 8; Indels 0; Gaps	
DB	2 ILTHKRLKLFHFGSLPSPFLYRTIDLGIGTF 37	
	:    :          : ::: :     :	
	192 ILTHKRLKLFMLFGKHLAPLEYLHKSPRELIGITW 227	
RESULT 8		
Q9W501		
ID	Q9W501 PRELIMINARY; PRT; 448 AA.	
AC	Q9W501;	
DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)	
GN	Hypothetical 51.5 kDa protein.	
GE	Y41G9A.3.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
QC	Rhabdiida; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
NP	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RX	MEDLINE=99069613; PubMed=9851916;	
RA	None;	
RT	"Genome sequence of the nematode C. elegans: a platform for	
KT	investigating biology. The C. elegans Sequencing Consortium.";	
RL	Science 282:2012-2018(1998).	
RP	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RA	Ryan E., Wohlman P., Walker C., Fielder T.;	
RT	"The sequence of C. elegans cosmid Y41G9A.";	
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.	
RP	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BRISTOL N2;	
RA	Waterston R.;	





Search completed: June 6, 2003, 11:02:53  
Job time : 7.31349 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2003, 10:59:34 ; Search time 2.44133 Seconds  
(without alignments)  
494.132 Million cell updates/sec

Title: US-09-826-581-6\_COPY\_320\_360

Perfect score: 213

Sequence: 1 HILYHKLLKFLHIFGSLP.....PSFLYRTIQDLGIGTRDLA 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	58.2	328	2	US-08-878-989-7
2	124	58.2	328	4	US-09-272-796-7
3	112	52.6	330	4	US-09-359-161-6
4	112	52.6	331	2	US-08-878-989-21
5	112	52.6	331	3	US-09-101-146-64
6	112	52.6	331	4	US-09-272-796-21
7	63	29.6	379	4	US-09-359-161-5
8	54.5	25.6	179	6	5463025-4
9	54.5	25.6	464	6	5463025-1
10	34	25.4	92	4	US-08-905-223-497
11	52.5	24.6	416	4	US-09-330-611-8
12	51	23.9	808	4	US-09-134-001C-3105
13	49	23.0	483	4	US-08-924-183-9
14	49	23.0	483	4	US-09-488-364-9
15	49	23.0	712	1	US-08-587-889-2
16	49	23.0	712	2	US-08-980-060-5
17	49	23.0	712	4	US-09-307-185-5
18	49	23.0	712	5	PCT-US96-09193-2
19	48	22.5	320	4	US-09-576-160B-4
20	48	22.5	350	4	US-09-576-160B-5
21	47.5	22.3	3289	2	US-08-477-451-2
22	47	22.1	314	4	US-09-188-930-193
23	47	22.1	316	4	US-09-188-930-337
24	46.5	21.8	352	4	US-09-576-160B-1
25	46.5	21.8	352	4	US-09-576-160B-2
26	46.5	21.8	352	4	US-09-576-160B-3
27	46.5	21.8	353	1	US-08-229-418-2

28	46.5	21.8	353	2	US-08-932-761A-2
29	46.5	21.8	353	4	US-09-307-912-2
30	46.5	21.8	353	5	PCT-US95-04464-2
31	46	21.6	260	4	US-09-134-001C-4507
32	46	21.6	471	4	US-09-330-611-10
33	46	21.6	652	1	US-08-261-663A-6
34	46	21.6	652	4	US-09-357-208A-5
35	46	21.6	652	5	PCT-US95-07754A-6
36	46	21.6	829	1	US-08-346-455B-34
37	46	21.6	829	3	US-08-977-221-34
38	46	21.6	829	4	US-09-483-831B-34
39	46	21.6	829	5	PCT-US95-06613-34
40	46	21.6	915	1	US-08-346-455B-69
41	46	21.6	915	3	US-08-977-221-69
42	46	21.6	915	4	US-09-483-831B-69
43	46	21.6	915	5	PCT-US95-06613-69
44	46	21.6	1143	2	US-08-310-912A-108
45	46	21.6	1143	4	US-09-301-085-108

## ALIGNMENTS

RESULT 1  
US-08-878-989-7  
; Sequence 7, Application US/08878989  
; Patent No. 5885803  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl G.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/878,989  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; APPLICATION DATA:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0321 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PENITUT01

Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4507, Ap  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 5, Appl  
Sequence 34, Appl  
Sequence 34, Appl  
Sequence 34, Appl  
Sequence 34, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 108, App  
Sequence 108, App

Qy	1	HILTHKRLLKFLHIFGSLLPPSFYRTIQDLGIGTFRDLA	41
		: :	
Db	156	YILTHKRLLKFLQFSDMPFAFKKONLDELGIGTYHNIA	196

ATTORNEY/AGENT INFORMATION

```
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1335856
US-08-878-989-21

Query Match 52.6%; Score 112; DB 2; Length 331;
Best Local Similarity 46.3%; Pred. No. 2.3e-08;
Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 HILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
Db 165 YILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTYANIA 205

RESULT 5
US-09-101-146-64
; Sequence 64, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; APPLICATION DATA:
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-101-146-64

Query Match 52.6%; Score 112; DB 3; Length 331;
Best Local Similarity 46.3%; Pred. No. 2.3e-08;
Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 HILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
Db 165 YILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTYANIA 205

RESULT 6
US-09-272-796-21
; Sequence 21, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1335856
US-09-272-796-21

Query Match 52.6%; Score 112; DB 4; Length 331;
Best Local Similarity 46.3%; Pred. No. 2.3e-08;
Matches 19; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 HILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTFRDLA 41
Db 165 YILTHKRLKLFHFGSLPRPSFLYRTIQDLGIGTYANIA 205

RESULT 7
US-09-359-161-5
; Sequence 5, Application US/09359161A
; Patent No. 6342656
; GENERAL INFORMATION:
; APPLICANT: Bradford, Kent J.
```

```
; APPLICANT: Dahal, Peetambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; TITLE OF INVENTION: to Stress Conditions in Plants
; FILE REFERENCE: 023070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
; FEATURE:
; OTHER INFORMATION: Phaseolus vulgaris L. Pv42
US-09-359-161-5

Query Match      29.6%; Score 63; DB 4; Length 379;
Best Local Similarity 48.6%; Pred. No. 0.29;
Matches 17; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

Qy      2 ILTHKRLKFLHIFGSLPRPSFLYRTIQDLGIGT 36
Db      186 MLTQMDLKLFLHGGGAE--HSILSRVQDLGADT 218

RESULT 8
5463025-4
; Patent No. 5463025
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 4
; LENGTH: 179
5463025-4

Query Match      25.6%; Score 54.5; DB 6; Length 179;
Best Local Similarity 48.6%; Pred. No. 2.1;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy      4 THKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFR 38
Db      129 TLQRLOQVLHA-GSGPCLPHLLSLRCLQDLGPGAFR 162

RESULT 9
5463025-1
; Patent No. 5463025
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
```

```
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 1
; LENGTH: 464
5463025-1

Query Match      25.6%; Score 54.5; DB 6; Length 464;
Best Local Similarity 48.6%; Pred. No. 6;
Matches 17; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy      4 THKRLKFLHIFGSLPRPSFLYRTIQDLGIGTFR 38
Db      101 TLQRLOQVLHA-GSGPCLPHLLSLRCLQDLGPGAFR 134

RESULT 10
US-08-905-223-497
; Sequence 497, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knodbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 497:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -26..-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 3.6
; OTHER INFORMATION: seq ARSLQLFLRLVGO/LK
US-08-905-223-497

Query Match      25.4%; Score 54; DB 4; Length 92;
Best Local Similarity 40.7%; Pred. No. 1.2;
Matches 11; Conservative 7; Mismatches 7; Indels 2; Gaps 1;

Qy      6 KRLKFLHIFGSL--LPRPSFLYRTIQ 30
Db      15 RSLQLFLRLVGO/LKVRPTGWYRNQ 41
```



```
RESULT 11
US-09-330-611-8
; Sequence 8, Application US/09330611
; Patent No. 6248874
; GENERAL INFORMATION:
; APPLICANT: FREY, Perry A.
; APPLICANT: RUZICKA, Frank J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING BACTERIAL LYSINE 2,3-AMINOMUTASE
; FILE REFERENCE: 032026/0476
; CURRENT APPLICATION NUMBER: US/09/330,611
; CURRENT FILING DATE: 1995-06-11
; EARLIER APPLICATION NUMBER: US 09/198,942
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-330-611-8

Query Match          24.6%; Score 52.5; DB 4; Length 416;
Best Local Similarity 40.5%; Pred. No. 10;
Matches 15; Conservative 5; Mismatches 12; Indels 5; Gaps 2;

QY      4 TH--KRLKFLHIFGSLPRPSFLYRTIQDLGIGTFR 38
      || ||| :| : || :| || ||| ||
DB      270 THVMKRL---VHLVKRMVRPYIYVCDLSLGIGHFR 303

RESULT 12
US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3105
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3105

Query Match          23.9%; Score 51; DB 4; Length 808;
Best Local Similarity 30.2%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 11; Indels 10; Gaps 2;

QY      2 ILTKRLKFLH-----IFGSLPRPSFLYRTIQDLGIGTGF 37
      || : : || || :| ||| : : ||
DB      1 ILKEINVSLHIAVLPLIFALIIP---FLYRVKRIHLGWF 40

RESULT 13
US-08-924-183-9
; Sequence 9, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
```

```
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-08-924-183-9

Query Match          23.0%; Score 49; DB 4; Length 483;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      21 RPSFLYRTIQDLGIGTFRDL 40
      :| ||| :| ||| :| :
DB      19 QPESLYRVVQTILGEGAFGEV 38

RESULT 14
US-09-488-364-9
; Sequence 9, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1013
; CURRENT APPLICATION NUMBER: US/09/488,364
; CURRENT FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-488-364-9

Query Match          23.0%; Score 49; DB 4; Length 483;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      21 RPSFLYRTIQDLGIGTFRDL 40
      :| ||| :| ||| :| :
DB      19 QPESLYRVVQTILGEGAFGEV 38

RESULT 15
US-08-587-889-2
; Sequence 2, Application US/08587889
; Patent No. 5654397
; GENERAL INFORMATION:
; APPLICANT: CAO, Zhaodan
; APPLICANT: CROSTON, Glenn E.
; APPLICANT: GOEDEL, David V.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,889
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-60916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-587-889-2

Query Match      23.0%; Score 49; DB 1; Length 712;
Best Local Similarity 37.1%; Pred. No. 59;
Matches 13; Conservative 3; Mismatches 7; Indels 12; Gaps 1;

QY      1 HILTHKRLKFLHIF-----GSLLPKPS 23
      |||||:|:| |
Db      85 HILTHQLLRARDIITAWHPAPLPSPCTTAPPS 119
```

Search completed: June 6, 2003, 11:04:15  
Job time : 3.44133 secs

Result No.	Query			Length	DB	ID	Description
	Score	Match	°				
1	213	100.0	489	10	US-09-826-581-6		Sequence 6, Appli
2	112	52.6	344	10	US-09-925-297-461		Sequence 461, App
3	105	49.3	1207	9	US-10-108-605-71		Sequence 71, Appl
4	54	25.4	284	9	US-09-510-332-47		Sequence 47, Appl
5	53	24.9	1023	9	US-10-028-072-200		Sequence 200, App
6	53	24.9	1023	9	US-10-121-049-200		Sequence 200, App
7	53	24.9	1023	9	US-10-123-904-200		Sequence 200, App
8	53	24.9	1023	9	US-10-140-470-200		Sequence 200, App
9	53	24.9	1023	9	US-10-175-746-200		Sequence 200, App
10	53	24.9	1023	9	US-10-176-918-200		Sequence 200, App
11	53	24.9	1023	9	US-10-176-921-200		Sequence 200, App
12	53	24.9	1023	9	US-10-137-865-200		Sequence 200, App
13	53	24.9	1023	9	US-10-140-474-200		Sequence 200, App
14	53	24.9	1023	9	US-10-142-431-200		Sequence 200, App
15	53	24.9	1023	9	US-10-143-114-200		Sequence 200, App
16	53	24.9	1023	9	US-10-140-002-200		Sequence 200, App
17	53	24.9	1023	9	US-10-142-419-200		Sequence 200, App
18	53	24.9	1023	9	US-10-123-262-200		Sequence 200, App
19	53	24.9	1023	9	US-10-142-423-200		Sequence 200, App

```

; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 47
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R26 (hGR26)
; NAME/KEY: MOD_RES
; LOCATION: (20)
; OTHER INFORMATION: xaa = any amino acid
; US-09-510-332-47

Best Match          25.4%; Score 54; DB 9; Length 264;
Query Local Similarity 21.9%; Pred. No. 10;
Matches 14; Conservative 11; Mismatches 11; Indels 28; Gaps 2

QY      1  HLTHKKLKLKFLHFGSL-----LPRSFLY-----RTIQDL 32
       1:  :|::|::|::|::|  |::|  |::|  |::|  |::|  |::|  |::|
Db      73  HLYAFNKLKFIHFWALTNLHVTWLACCLSVFYFFKIAFYSHPCFIWLRWIRSLTLEL 132
       QY      33  GIGT 36
       Db      133  PLGS 136
               :|:
               :|:

RESULT 5
US-10-028-072-200
; Sequence 200, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24

```

1	PRIOR APPLICATION NUMBER: 60/062250	2	PRIOR FILING DATE: 1998-02-27
2	PRIOR FILING DATE: 1997-10-17	3	PRIOR APPLICATION NUMBER: 60/079728
3	PRIOR APPLICATION NUMBER: 60/062285	4	PRIOR FILING DATE: 1998-03-27
4	PRIOR FILING DATE: 1997-10-17	5	PRIOR APPLICATION NUMBER: 60/080165
5	PRIOR APPLICATION NUMBER: 60/062287	6	PRIOR FILING DATE: 1998-03-31
6	PRIOR FILING DATE: 1997-10-17	7	PRIOR APPLICATION NUMBER: 60/081203
7	PRIOR APPLICATION NUMBER: 60/062814	8	PRIOR FILING DATE: 1998-04-09
8	PRIOR FILING DATE: 1997-10-24	9	PRIOR APPLICATION NUMBER: 60/081229
9	PRIOR APPLICATION NUMBER: 60/062816	10	PRIOR FILING DATE: 1998-04-09
10	PRIOR FILING DATE: 1997-10-24	11	PRIOR APPLICATION NUMBER: 60/081695
11	PRIOR APPLICATION NUMBER: 60/063045	12	PRIOR FILING DATE: 1998-04-14
12	PRIOR FILING DATE: 1997-10-24	13	PRIOR APPLICATION NUMBER: 60/081817
13	PRIOR APPLICATION NUMBER: 60/063082	14	PRIOR FILING DATE: 1998-04-15
14	PRIOR FILING DATE: 1997-10-31	15	PRIOR APPLICATION NUMBER: 60/081818
15	PRIOR APPLICATION NUMBER: 60/063127	16	PRIOR FILING DATE: 1998-04-15
16	PRIOR FILING DATE: 1997-10-24	17	PRIOR APPLICATION NUMBER: 60/082999
17	PRIOR APPLICATION NUMBER: 60/063327	18	PRIOR FILING DATE: 1998-04-24
18	PRIOR FILING DATE: 1997-10-27	19	PRIOR APPLICATION NUMBER: 60/083322
19	PRIOR APPLICATION NUMBER: 60/063329	20	PRIOR FILING DATE: 1998-04-28
20	PRIOR FILING DATE: 1997-10-27	21	PRIOR APPLICATION NUMBER: 60/083545
21	PRIOR APPLICATION NUMBER: 60/063550	22	PRIOR FILING DATE: 1998-04-29
22	PRIOR FILING DATE: 1997-10-28	23	PRIOR APPLICATION NUMBER: 60/084600
23	PRIOR APPLICATION NUMBER: 60/063561	24	PRIOR FILING DATE: 1998-05-07
24	PRIOR FILING DATE: 1997-10-28	25	PRIOR APPLICATION NUMBER: 60/084627
25	PRIOR APPLICATION NUMBER: 60/063704	26	PRIOR FILING DATE: 1998-05-07
26	PRIOR FILING DATE: 1997-10-29	27	PRIOR APPLICATION NUMBER: 60/084637
27	PRIOR APPLICATION NUMBER: 60/063733	28	PRIOR FILING DATE: 1998-05-07
28	PRIOR FILING DATE: 1997-10-29	29	PRIOR APPLICATION NUMBER: 60/085149
29	PRIOR APPLICATION NUMBER: 60/063735	30	PRIOR FILING DATE: 1998-05-12
30	PRIOR FILING DATE: 1997-10-29	31	PRIOR APPLICATION NUMBER: 60/085323
31	PRIOR APPLICATION NUMBER: 60/063738	32	PRIOR FILING DATE: 1998-05-13
32	PRIOR FILING DATE: 1997-10-29	33	PRIOR APPLICATION NUMBER: 60/085338
33	PRIOR APPLICATION NUMBER: 60/063755	34	PRIOR FILING DATE: 1998-05-13
34	PRIOR FILING DATE: 1997-10-17	35	PRIOR APPLICATION NUMBER: 60/085339
35	PRIOR APPLICATION NUMBER: 60/064248	36	PRIOR FILING DATE: 1998-05-13
36	PRIOR FILING DATE: 1997-11-03	37	PRIOR APPLICATION NUMBER: 60/085579
37	PRIOR APPLICATION NUMBER: 60/064809	38	PRIOR FILING DATE: 1998-05-15
38	PRIOR FILING DATE: 1997-11-07	39	PRIOR APPLICATION NUMBER: 60/085697
39	PRIOR APPLICATION NUMBER: 60/065186	40	PRIOR FILING DATE: 1998-05-15
40	PRIOR FILING DATE: 1997-11-12	41	PRIOR APPLICATION NUMBER: 60/085704
41	PRIOR APPLICATION NUMBER: 60/065846	42	PRIOR FILING DATE: 1998-05-15
42	PRIOR FILING DATE: 1997-11-17	43	PRIOR APPLICATION NUMBER: 60/086414
43	PRIOR APPLICATION NUMBER: 60/066364	44	PRIOR FILING DATE: 1998-05-22
44	PRIOR FILING DATE: 1997-11-21	45	PRIOR APPLICATION NUMBER: 60/086430
45	PRIOR APPLICATION NUMBER: 60/066453	46	PRIOR FILING DATE: 1998-05-22
46	PRIOR FILING DATE: 1997-11-24	47	PRIOR APPLICATION NUMBER: 60/087106
47	PRIOR APPLICATION NUMBER: 60/066511	48	PRIOR FILING DATE: 1998-05-28
48	PRIOR FILING DATE: 1997-11-24	49	PRIOR APPLICATION NUMBER: 60/088026
49	PRIOR APPLICATION NUMBER: 60/066770	50	PRIOR FILING DATE: 1998-06-04
50	PRIOR FILING DATE: 1997-11-24	51	PRIOR APPLICATION NUMBER: 60/088730
51	PRIOR APPLICATION NUMBER: 60/069212	52	PRIOR FILING DATE: 1998-06-10
52	PRIOR FILING DATE: 1997-12-11	53	PRIOR APPLICATION NUMBER: 60/088741
53	PRIOR APPLICATION NUMBER: 60/069278	54	PRIOR FILING DATE: 1998-06-10
54	PRIOR FILING DATE: 1997-12-11	55	PRIOR APPLICATION NUMBER: 60/088810
55	PRIOR APPLICATION NUMBER: 60/069334	56	PRIOR FILING DATE: 1998-06-10
56	PRIOR FILING DATE: 1997-12-11	57	PRIOR APPLICATION NUMBER: 60/088858
57	PRIOR APPLICATION NUMBER: 60/069694	58	PRIOR FILING DATE: 19/98-06-11
58	PRIOR FILING DATE: 1997-12-16	59	PRIOR APPLICATION NUMBER: 60/089532
59	PRIOR APPLICATION NUMBER: 60/072320	60	PRIOR FILING DATE: 1998-06-17
60	PRIOR FILING DATE: 1998-01-23	61	PRIOR APPLICATION NUMBER: 60/089599
61	PRIOR APPLICATION NUMBER: 60/073612	62	PRIOR FILING DATE: 1998-06-17
62	PRIOR FILING DATE: 1998-02-04	63	PRIOR APPLICATION NUMBER: 60/089907
63	PRIOR APPLICATION NUMBER: 60/074086	64	PRIOR FILING DATE: 1998-06-18
64	PRIOR FILING DATE: 1998-02-09	65	PRIOR APPLICATION NUMBER: 60/089947
65	PRIOR APPLICATION NUMBER: 60/074092	66	PRIOR FILING DATE: 1998-06-19
66	PRIOR FILING DATE: 1998-02-09	67	PRIOR APPLICATION NUMBER: 60/090349
67	PRIOR APPLICATION NUMBER: 60/077791	68	PRIOR FILING DATE: 1998-06-23
68	PRIOR FILING DATE: 1998-03-12	69	PRIOR APPLICATION NUMBER: 60/090429
69	PRIOR APPLICATION NUMBER: 60/078910		





```
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKFLHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGOVIRSPTNTISVYRTFQDDGLGTFQ 383

RESULT 13
US-10-140-474-200
; Sequence 200, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;
```

```
QY      2 ILTHKRLKFLHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGOVIRSPTNTISVYRTFQDDGLGTFQ 383

RESULT 14
US-10-142-431-200
; Sequence 200, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;

QY      2 ILTHKRLKFLHFGSLPRP-----SFLYRTIQDLGIGTFR 38
Db      347 VLANQTLL-----VEGOVIRSPTNTISVYRTFQDDGLGTFQ 383

RESULT 15
US-10-143-114-200
; Sequence 200, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 200
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-114-200

Query Match      24.9%; Score 53; DB 9; Length 1023;
Best Local Similarity 34.1%; Pred. No. 66;
Matches 14; Conservative 9; Mismatches 10; Indels 8; Gaps 2;
```





US 098265810NP1



Creation date: 08-07-2003

Indexing Officer: PRT1634 - GAU1634 PRINTER

Team: GAU1634PrtWorkingFolder

Dossier: 09826581

Legal Date: 16-06-2003

No.	Doccode	Number of pages
1	SRNT	562

Total number of pages: 562

Remarks:

Order of re-scan issued on .....